GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 10:10:38 ; Search time 692 Seconds
(without alignments)
587.535 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 8765912 seqs, 3504951483 residues

Word size: 1

Total number of hits satisfying chosen parameters: 10478022

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135845_19714/app_query.fasta_1.199
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n_olig_sz30.rnpb
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30
-USER=US10031158_@CGN_1_1_723_@runat_14102005_135845_19714 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	7	12.1	23	9	US-09-860-996-14	Sequence 14, Appl
	2	7	12.1	25	20	US-10-483-654-9	Sequence 9, Appli
	3	7	12.1	25	21	US-10-719-900-27595	Sequence 27595, A
c	4	7	12.1	25	24	US-11-036-317-217607	Sequence 217607,
c	5	7	12.1	25	24	US-11-036-317-236358	Sequence 236358,
c	6	7	12.1	25	24	US-11-036-317-358635	Sequence 358635,
c	7	7	12.1	25	24	US-11-036-317-390331	Sequence 390331,
c	8	7	12.1	25	24	US-11-036-317-397073	Sequence 397073,
c	9	7	12.1	25	24	US-11-036-317-898261	Sequence 898261,
c	10	7	12.1	25	24	US-11-036-317-968427	Sequence 968427,
c	11	7	12.1	25	24	US-11-036-317-989913	Sequence 989913,
c	12	6	10.3	19	18	US-10-092-900A-680	Sequence 680, App
	13	6	10.3	20	17	US-10-181-874-18	Sequence 18, Appl
	14	6	10.3	20	17	US-10-271-602B-20	Sequence 20, Appl
	15	6	10.3	20	17	US-10-271-602B-29	Sequence 29, Appl
	16	6	10.3	20	17	US-10-271-602B-50	Sequence 50, Appl
	17	6	10.3	20	17	US-10-349-143-11305	Sequence 11305, A
c	18	6	10.3	20	20	US-10-473-126-468	Sequence 468, App
	19	6	10.3	21	20	US-10-751-736-32850	Sequence 32850, A
	20	6	10.3	21	21	US-10-847-918-13631	Sequence 13631, A
	21	6	10.3	21	21	US-10-847-918-13633	Sequence 13633, A
	22	6	10.3	21	21	US-10-847-918-13634	Sequence 13634, A
c	23	6	10.3	21	21	US-10-847-918-13635	Sequence 13635, A
	24	6	10.3	21	21	US-10-847-918-13639	Sequence 13639, A
	25	6	10.3	21	21	US-10-847-918-13640	Sequence 13640, A
c	26	6	10.3	21	21	US-10-847-918-13641	Sequence 13641, A
	27	6	10.3	21	21	US-10-847-918-13642	Sequence 13642, A
c	28	6	10.3	21	21	US-10-847-918-13644	Sequence 13644, A
c	29	6	10.3	22	16	US-10-302-279-20	Sequence 0, Appli
	30	6	10.3	23	9	US-09-942-374-6	Sequence 6, Appli
	31	6	10.3	23	18	US-10-665-956-6	Sequence 6, Appli
	32	6	10.3	23	20	US-10-483-654-8	Sequence 8, Appli
	33	6	10.3	25	15	US-10-098-263B-31202	Sequence 31202, A
c	34	6	10.3	25	15	US-10-098-263B-67759	Sequence 67759, A
	35	6	10.3	25	15	US-10-098-263B-78553	Sequence 78553, A
	36	6	10.3	25	15	US-10-098-263B-78554	Sequence 78554, A
c	37	6	10.3	25	15	US-10-098-263B-83546	Sequence 83546, A
c	38	6	10.3	25	15	US-10-098-263B-121122	Sequence 121122,
c	39	6	10.3	25	21	US-10-719-900-45610	Sequence 45610, A
	40	6	10.3	25	21	US-10-719-900-46685	Sequence 46685, A
	41	6	10.3	25	21	US-10-719-900-46686	Sequence 46686, A
	42	6	10.3	25	21	US-10-719-900-47953	Sequence 47953, A
c	43	6	10.3	25	21	US-10-719-900-124417	Sequence 124417,
c	44	6	10.3	25	21	US-10-719-900-174340	Sequence 174340,
	45	6	10.3	25	21	US-10-719-900-241517	Sequence 241517,

ALIGNMENTS

RESULT 1
US-09-860-996-14/c
; Sequence 14, Application US/09860996
; Patent No. US20020034393A1
; GENERAL INFORMATION:
; APPLICANT: Mitrophanous, et al
; TITLE OF INVENTION: VECTOR
; FILE REFERENCE: 674523-2010
; CURRENT APPLICATION NUMBER: US/09/860,996
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/GB99/03866
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 9825524.3
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 14
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Human Immunodeficiency Virus Type 1
US-09-860-996-14

Alignment Scores:
Pred. No.:              107          Length:        23
Score:                  7.00         Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            12.07%       Indels:        0
DB:                     9            Gaps:          0

US-10-031-158-14 (1-58) x US-09-860-996-14 (1-23)

Qy          5 ProProSerProLeuPhePhe 11
              |||||||||||||||||||||
Db         22 CCCCCATCCCCCCTTTTCTTT 2

RESULT 2
US-10-483-654-9
; Sequence 9, Application US/10483654
; Publication No. US20040220095A1
; GENERAL INFORMATION:
;  APPLICANT: Braun, Klaus
;  APPLICANT:  Waldeck, Waldemar
;  APPLICANT:  Pipkorn, Rudiger
;  APPLICANT:  Debus, Jurgen
;  APPLICANT:  Braun, Isabell
;  TITLE OF INVENTION: PNA Conjugate for the Treatment of Diseases Associated with HIV
;  FILE REFERENCE: 4121-159
;  CURRENT APPLICATION NUMBER: US/10/483,654
;  CURRENT FILING DATE:  2004-01-12
;  PRIOR APPLICATION NUMBER: PCT/DE 02/02564
;  PRIOR FILING DATE: 2002-07-12
;  PRIOR APPLICATION NUMBER: DE 101 33 307.2
;  PRIOR FILING DATE: 2001-07-12
;  NUMBER OF SEQ ID NOS: 54
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PNA Ia
US-10-483-654-9

Alignment Scores:
Pred. No.:              115          Length:        25
Score:                  7.00         Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            12.07%       Indels:        0
DB:                     20           Gaps:          0

US-10-031-158-14 (1-58) x US-10-483-654-9 (1-25)

Qy          7 SerProLeuPhePhePheLeu 13
              |||||||||||||||||||||
Db          4 TCCCCCCTTTTCTTTTTTCTT 24

RESULT 3
US-10-719-900-27595
; Sequence 27595, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
;  APPLICANT: Xue Mei Zhou
;  TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
;  FILE REFERENCE: 3528.1
;  CURRENT APPLICATION NUMBER: US/10/719,900
;  CURRENT FILING DATE:  2003-11-20
;  PRIOR APPLICATION NUMBER: 60/427,808
;  PRIOR FILING DATE: 2002 11 20
;  NUMBER OF SEQ ID NOS: 982914
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 27595
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-10-719-900-27595

Alignment Scores:
Pred. No.:              115          Length:        25
Score:                  7.00         Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            12.07%       Indels:        0
DB:                     21           Gaps:          0

US-10-031-158-14 (1-58) x US-10-719-900-27595 (1-25)

Qy         35 LeuMetLeuLeuArgTyrIle 41
              |||||||||||||||||||||
Db          5 CTAATGTTACTCAGGTATATA 25

RESULT 4
US-11-036-317-217607/c
; Sequence 217607, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Alan
;  APPLICANT:  Blume, John
;  TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
;  FILE REFERENCE: 3654.1
;  CURRENT APPLICATION NUMBER: US/11/036,317
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: US 60/536,639
;  PRIOR FILING DATE: 2004-01-13
;  NUMBER OF SEQ ID NOS: 991174
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 217607
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-11-036-317-217607

Alignment Scores:
Pred. No.:              115          Length:        25
Score:                  7.00         Matches:       7
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            12.07%       Indels:        0
DB:                     24           Gaps:          0

US-10-031-158-14 (1-58) x US-11-036-317-217607 (1-25)

Qy         32 AsnPheSerLeuMetLeuLeu 38
              ||||||||||||||||||||||
Db         21 AATTTTTCTTTGATGTTGTTG 1

RESULT 5
US-11-036-317-236358/c
; Sequence 236358, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Alan
;  APPLICANT:  Blume, John
;  TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
;  FILE REFERENCE: 3654.1
;  CURRENT APPLICATION NUMBER: US/11/036,317
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: US 60/536,639
;  PRIOR FILING DATE: 2004-01-13
;  NUMBER OF SEQ ID NOS: 991174
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
```

; SEQ ID NO 236358
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-236358

Alignment Scores:

Pred. No.:	115	Length:	25
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	24	Gaps:	0

US-10-031-158-14 (1-58) x US-11-036-317-236358 (1-25)

```
Qy         32 AsnPheSerLeuMetLeuLeu 38
              ||||||||||||||||||||||
Db         22 AATTTTTCTTTGATGTTGTTG 2
```

RESULT 6
US-11-036-317-358635/c
; Sequence 358635, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 358635
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-358635

Alignment Scores:

Pred. No.:	115	Length:	25
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	24	Gaps:	0

US-10-031-158-14 (1-58) x US-11-036-317-358635 (1-25)

```
Qy         32 AsnPheSerLeuMetLeuLeu 38
              ||||||||||||||||||||||
Db         23 AATTTTTCTTTGATGTTGTTG 3
```

RESULT 7
US-11-036-317-390331/c
; Sequence 390331, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 390331
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-390331

Alignment Scores:

Pred. No.:	115	Length:	25
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	24	Gaps:	0

US-10-031-158-14 (1-58) x US-11-036-317-390331 (1-25)

```
Qy         32 AsnPheSerLeuMetLeuLeu 38
              ||||||||||||||||||||||
Db         24 AATTTTTCTTTGATGTTGTTG 4
```

RESULT 8
US-11-036-317-397073/c
; Sequence 397073, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 397073
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-397073

Alignment Scores:

Pred. No.:	115	Length:	25
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	24	Gaps:	0

US-10-031-158-14 (1-58) x US-11-036-317-397073 (1-25)

```
Qy         32 AsnPheSerLeuMetLeuLeu 38
              ||||||||||||||||||||||
Db         25 AATTTTTCTTTGATGTTGTTG 5
```

RESULT 9
US-11-036-317-898261/c
; Sequence 898261, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Alan
; APPLICANT: Blume, John
; TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
; FILE REFERENCE: 3654.1
; CURRENT APPLICATION NUMBER: US/11/036,317
; CURRENT FILING DATE: 2005-01-13
; PRIOR APPLICATION NUMBER: US 60/536,639
; PRIOR FILING DATE: 2004-01-13
; NUMBER OF SEQ ID NOS: 991174
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 898261
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-036-317-898261

```
Alignment Scores:
Pred. No.:                115             Length:         25
Score:                    7.00            Matches:        7
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              12.07%          Indels:         0
DB:                       24              Gaps:           0

US-10-031-158-14 (1-58) x US-11-036-317-898261 (1-25)

Qy          32 AsnPheSerLeuMetLeuLeu 38
               ||||||||||||||||||||||
Db          22 AATTTTTCTTTGATGTTGTTG 2

RESULT 10
US-11-036-317-968427/c
; Sequence 968427, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Alan
;  APPLICANT:  Blume, John
;  TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
;  FILE REFERENCE: 3654.1
;  CURRENT APPLICATION NUMBER: US/11/036,317
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: US 60/536,639
;  PRIOR FILING DATE: 2004-01-13
;  NUMBER OF SEQ ID NOS: 991174
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 968427
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-11-036-317-968427

Alignment Scores:
Pred. No.:                115             Length:         25
Score:                    7.00            Matches:        7
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              12.07%          Indels:         0
DB:                       24              Gaps:           0

US-10-031-158-14 (1-58) x US-11-036-317-968427 (1-25)

Qy          32 AsnPheSerLeuMetLeuLeu 38
               ||||||||||||||||||||||
Db          23 AATTTTTCTTTGATGTTGTTG 3

RESULT 11
US-11-036-317-989913/c
; Sequence 989913, Application US/11036317
; Publication No. US20050214823A1
; GENERAL INFORMATION:
;  APPLICANT: Williams, Alan
;  APPLICANT:  Blume, John
;  TITLE OF INVENTION: Method of Analysis of Alternative Splicing in Mouse
;  FILE REFERENCE: 3654.1
;  CURRENT APPLICATION NUMBER: US/11/036,317
;  CURRENT FILING DATE:  2005-01-13
;  PRIOR APPLICATION NUMBER: US 60/536,639
;  PRIOR FILING DATE: 2004-01-13
;  NUMBER OF SEQ ID NOS: 991174
;  SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 989913
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-11-036-317-989913

Alignment Scores:
Pred. No.:                115             Length:         25
Score:                    7.00            Matches:        7
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              12.07%          Indels:         0
DB:                       24              Gaps:           0

US-10-031-158-14 (1-58) x US-11-036-317-989913 (1-25)

Qy          32 AsnPheSerLeuMetLeuLeu 38
               ||||||||||||||||||||||
Db          25 AATTTTTCTTTGATGTTGTTG 5

RESULT 12
US-10-092-900A-680/c
; Sequence 680, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
;  APPLICANT: Padigaru, Muralidhara
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Shenoy, Suresh G.
;  APPLICANT:  Taupier Jr., Raymond J.
;  APPLICANT:  Pena, Carol E.A.
;  APPLICANT:  Li, Li
;  APPLICANT:  Zerhusen, Bryan D.
;  APPLICANT:  Gusev, Vladimir Y.
;  APPLICANT:  Ji, Weizhen
;  APPLICANT:  Gorman, Linda
;  APPLICANT:  Miller, Charles E.
;  APPLICANT:  Kekuda, Ramesh
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Gangolli, Esha A.
;  APPLICANT:  Vernet, Corine A.M.
;  APPLICANT:  Guo, Xiaojia Sasha
;  APPLICANT:  Tchernev,, Velizar T.
;  APPLICANT:  Fernandes, Elma R.
;  APPLICANT:  Casman, Stacie J.
;  APPLICANT:  Malyankar, Uriel M.
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  Liu, Yi
;  APPLICANT:  Anderson, David W.
;  APPLICANT:  Spaderna, Steven K.
;  APPLICANT:  Catterton, Elina
;  APPLICANT:  Leite, Mario W.
;  APPLICANT:  Zhong, Haihong
;  APPLICANT:  Alsobrook, John P.
;  APPLICANT:  Lepley, Denise M.
;  APPLICANT:  Rieger, Daniel K.
;  APPLICANT:  Burgess, Catherine E.
;  TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
;  FILE REFERENCE: 21402-290C
;  CURRENT APPLICATION NUMBER: US/10/092,900A
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: USSN 60/274,322
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/283,675
;  PRIOR FILING DATE: 2001-04-13
;  PRIOR APPLICATION NUMBER: USSN 60/338,092
;  PRIOR FILING DATE: 2001-12-03
;  PRIOR APPLICATION NUMBER: USSN 60/274,281
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/274,191
;  PRIOR FILING DATE: 2001-03-08
;  PRIOR APPLICATION NUMBER: USSN 60/325,681
;  PRIOR FILING DATE: 2001-09-27
;  PRIOR APPLICATION NUMBER: USSN 60/304,354
;  PRIOR FILING DATE: 2001-07-10
;  PRIOR APPLICATION NUMBER: USSN 60/279,995
;  PRIOR FILING DATE: 2001-03-30
;  PRIOR APPLICATION NUMBER: USSN 60/294,899
;  PRIOR FILING DATE: 2001-05-31
;  PRIOR APPLICATION NUMBER: USSN 60/287,424
;  PRIOR FILING DATE: 2001-04-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 768
```

; SEQ ID NO 680
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Forward Primer
US-10-092-900A-680

Alignment Scores:

Pred. No.:	934	Length:	19
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	18	Gaps:	0

US-10-031-158-14 (1-58) x US-10-092-900A-680 (1-19)

```
Qy         48 ThrArgPheTrpAspPro 53
              |||||||||||||||||
Db         19 ACTAGATTCTGGGACCCG 2
```

RESULT 13
US-10-181-874-18
; Sequence 18, Application US/10181874
; Publication No. US20030212020A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Susan Murray
; APPLICANT: Lex M. Cowsert
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF MACROPHAGE MIGRATION INHIBITORY FACTOR E
; FILE REFERENCE: RTSP-0351
; CURRENT APPLICATION NUMBER: US/10/181,874
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/489,869
; PRIOR FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 18
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-181-874-18

Alignment Scores:

Pred. No.:	978	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	17	Gaps:	0

US-10-031-158-14 (1-58) x US-10-181-874-18 (1-20)

```
Qy         52 AspProArgArgGlyThr 57
              ||||||||||||||||||
Db          3 GACCCGCGCAGAGGCACA 20
```

RESULT 14
US-10-271-602B-20
; Sequence 20, Application US/10271602B
; Publication No. US20040002073A1
; GENERAL INFORMATION:
; APPLICANT: Alice Xiang Li
; APPLICANT: Ghazala Hashmi
; APPLICANT: Michael Seul
; TITLE OF INVENTION: MULTIPLEXED ANALYSIS OF POLYMORPHIC LOCI
; TITLE OF INVENTION: BY CONCURRENT INTERROGATION AND ENZYME-MEDIATED DETECTION
; FILE REFERENCE: eMAP-US
; CURRENT APPLICATION NUMBER: US/10/271,602B
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/329,427
; PRIOR FILING DATE: 2001-10-14
; PRIOR APPLICATION NUMBER: 60/329,620
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/329,428
; PRIOR FILING DATE: 2001-10-14
; PRIOR APPLICATION NUMBER: 60/329,619
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/364,416
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Probe sequence derived from human genomic sequence
US-10-271-602B-20

Alignment Scores:

Pred. No.:	978	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	17	Gaps:	0

US-10-031-158-14 (1-58) x US-10-271-602B-20 (1-20)

```
Qy         12 PheLeuGlnLeuLeuLys 17
              ||||||||||||||||||
Db          1 TTCTTGCAGCTGCTTAAG 18
```

RESULT 15
US-10-271-602B-29
; Sequence 29, Application US/10271602B
; Publication No. US20040002073A1
; GENERAL INFORMATION:
; APPLICANT: Alice Xiang Li
; APPLICANT: Ghazala Hashmi
; APPLICANT: Michael Seul
; TITLE OF INVENTION: MULTIPLEXED ANALYSIS OF POLYMORPHIC LOCI
; TITLE OF INVENTION: BY CONCURRENT INTERROGATION AND ENZYME-MEDIATED DETECTION
; FILE REFERENCE: eMAP-US
; CURRENT APPLICATION NUMBER: US/10/271,602B
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/329,427
; PRIOR FILING DATE: 2001-10-14
; PRIOR APPLICATION NUMBER: 60/329,620
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/329,428
; PRIOR FILING DATE: 2001-10-14
; PRIOR APPLICATION NUMBER: 60/329,619
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/364,416
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Human
US-10-271-602B-29

Alignment Scores:

Pred. No.:	978	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	17	Gaps:	0

US-10-031-158-14 (1-58) x US-10-271-602B-29 (1-20)

```
Qy          12 PheLeuGlnLeuLeuLys 17
               ||||||||||||||||||
Db           1 TTCTTGCAGCTGCTTAAG 18
```

Search completed: October 16, 2005, 12:50:52
Job time : 693 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 09:22:02 ; Search time 130 Seconds
(without alignments)
730.031 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Word size: 1

Total number of hits satisfying chosen parameters: 965324

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135843_19628/app_query.fasta_1.199
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n_olig_sz30.rni -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30
-USER=US10031158_@CGN_1_1_105_@runat_14102005_135843_19628 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		6	10.3	20	3	US-09-489-869-18	Sequence 18, Appl
2		6	10.3	20	4	US-09-422-978-11305	Sequence 11305, A
3	c	6	10.3	22	4	US-08-857-636-20	Sequence 20, Appl
4	c	6	10.3	25	4	US-09-396-196G-41485	Sequence 41485, A
5	c	6	10.3	25	4	US-09-396-196G-41486	Sequence 41486, A
6	c	6	10.3	25	4	US-09-396-196G-41487	Sequence 41487, A
7	c	6	10.3	25	4	US-09-396-196G-41488	Sequence 41488, A
8	c	6	10.3	25	4	US-09-396-196G-41489	Sequence 41489, A
9	c	6	10.3	25	4	US-09-396-196G-41490	Sequence 41490, A
10	c	6	10.3	25	4	US-09-396-196G-113910	Sequence 113910,
11		6	10.3	25	4	US-09-396-196G-127372	Sequence 127372,
12		6	10.3	25	4	US-09-396-196G-127373	Sequence 127373,
13		6	10.3	25	4	US-09-396-196G-127384	Sequence 127384,
14	c	6	10.3	30	1	US-08-186-229-20	Sequence 20, Appl
15	c	6	10.3	30	2	US-08-470-124-20	Sequence 20, Appl
16	c	5	8.6	16	1	US-08-087-387-1	Sequence 1, Appli
17	c	5	8.6	16	1	US-08-455-627-1	Sequence 1, Appli
18	c	5	8.6	16	1	US-08-242-664-21	Sequence 21, Appl
19	c	5	8.6	16	1	US-08-242-664-22	Sequence 22, Appl
20	c	5	8.6	16	1	US-08-484-138-21	Sequence 21, Appl
21	c	5	8.6	16	1	US-08-484-138-22	Sequence 22, Appl
22	c	5	8.6	16	1	US-08-461-271-1	Sequence 1, Appli
23	c	5	8.6	16	1	US-08-713-685A-1	Sequence 1, Appli
24	c	5	8.6	16	2	US-08-689-856-1	Sequence 1, Appli
25	c	5	8.6	16	2	US-08-412-376-1	Sequence 1, Appli
26	c	5	8.6	16	3	US-09-070-477-1	Sequence 1, Appli
27	c	5	8.6	16	3	US-07-808-452-17	Sequence 17, Appl
28	c	5	8.6	16	3	US-09-133-717-11	Sequence 11, Appl
29	c	5	8.6	16	3	US-08-787-321-1	Sequence 1, Appli
30		5	8.6	16	3	US-08-787-321-16	Sequence 16, Appl
31	c	5	8.6	16	3	US-09-158-863C-11	Sequence 11, Appl
32	c	5	8.6	16	5	PCT-US92-10770-17	Sequence 17, Appl
33	c	5	8.6	16	5	PCT-US92-10792-15	Sequence 15, Appl
34	c	5	8.6	16	5	PCT-US95-06379-21	Sequence 21, Appl
35	c	5	8.6	16	5	PCT-US95-06379-22	Sequence 22, Appl
36	c	5	8.6	17	1	US-08-105-483-85	Sequence 85, Appl
37	c	5	8.6	17	1	US-08-709-209-85	Sequence 85, Appl
38	c	5	8.6	17	1	US-08-458-101-85	Sequence 85, Appl
39	c	5	8.6	17	3	US-09-049-277B-17	Sequence 17, Appl
40		5	8.6	17	3	US-08-584-040-2777	Sequence 2777, Ap
41		5	8.6	17	3	US-08-584-040-2778	Sequence 2778, Ap
42		5	8.6	17	3	US-08-584-040-2779	Sequence 2779, Ap
43		5	8.6	17	3	US-08-679-645-850	Sequence 850, App
44		5	8.6	17	3	US-08-679-645-852	Sequence 852, App
45		5	8.6	17	4	US-09-371-772B-1301	Sequence 1301, Ap

ALIGNMENTS

RESULT 1
US-09-489-869-18
; Sequence 18, Application US/09489869A
; Patent No. 6268151
; GENERAL INFORMATION:
; APPLICANT: Susan Murray
; APPLICANT: Lex M. Cowsert
; APPLICANT: Jacqueline Wyatt
; TITLE OF INVENTION: ANTISENSE MODULATION OF MACROPHAGE MIGRATION INHIBITORY FACTOR
; FILE REFERENCE: RTS-0110
; CURRENT APPLICATION NUMBER: US/09/489,869A
; CURRENT FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 88
; SEQ ID NO 18
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-489-869-18

Alignment Scores:

Pred. No.:	178	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	3	Gaps:	0

US-10-031-158-14 (1-58) x US-09-489-869-18 (1-20)

```
Qy      52 AspProArgArgGlyThr 57
           ||||||||||||||||||
Db       3 GACCCGCGCAGAGGCACA 20
```

RESULT 2

```
US-09-422-978-11305
; Sequence 11305, Application US/09422978
; Patent No. 6537751
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT: Blumenfeld, Marta
;  APPLICANT: Chumakov, Ilya
;  TITLE OF INVENTION: Biallelic markers for use in constructing a high density...
;  FILE REFERENCE: GENSET.020CP1
;  CURRENT APPLICATION NUMBER: US/09/422,978
;  CURRENT FILING DATE: 1999-10-20
;  EARLIER APPLICATION NUMBER: US 09/298,850
;  EARLIER FILING DATE: 1999-04-21
;  EARLIER APPLICATION NUMBER: US 60/109,732
;  EARLIER FILING DATE: 1998-11-23
;  EARLIER APPLICATION NUMBER: US 60/082,614
;  EARLIER FILING DATE: 1998-04-21
;  NUMBER OF SEQ ID NOS: 11796
; SEQ ID NO 11305
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: 1..20
;   OTHER INFORMATION: downstream amplification primer 99-4119 for SEQ 3440, in compleme
US-09-422-978-11305

Alignment Scores:
Pred. No.:                178             Length:         20
Score:                    6.00            Matches:        6
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              10.34%          Indels:         0
DB:                       4               Gaps:           0

US-10-031-158-14 (1-58) x US-09-422-978-11305 (1-20)

Qy          8 ProLeuPhePhePheLeu 13
              ||||||||||||||||||
Db          3 CCCCTATTTTTTTTTCTG 20

RESULT 3
US-08-857-636-20/c
; Sequence 20, Application US/08857636
; Patent No. 6552181
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Michael Carlton
;    APPLICANT:  Hahn, Heidi Eve
;    APPLICANT:  Wicking, Carol
;    APPLICANT:  Christiansen, Jeffrey
;    APPLICANT:  Zaphiropoulos, Peter G.
;    APPLICANT:  Gailani, Mae R.
;    APPLICANT:  Shanley, Susan Mary
;    APPLICANT:  Chidambaram, Abirami
;    APPLICANT:  Vorechovsky, Igor
;    APPLICANT:  Holmberg-Lindstrom, Erika
;    APPLICANT:  Unden, Anne Birgitte
;    APPLICANT:  Gillies, Susan Alana
;    APPLICANT:  Negus, Kylie
;    APPLICANT:  Smyth, Ian Mcleod
;    APPLICANT:  Pressman, Carol Leah
;    APPLICANT:  Leffell, David J.
;    APPLICANT:  Gerrard, Bernard
;    APPLICANT:  Goldstein, Alisa Miriam
;    APPLICANT:  Wainwright, Brandon
;    APPLICANT:  Toftgard, Rune Carl-Magnus
;    APPLICANT:  Chenevix-Trench, Georgia
;    APPLICANT:  Bale, Allen E.
;    TITLE OF INVENTION:  A Basal Cell Carcinoma Tumor Suppressor Gene
;    NUMBER OF SEQUENCES:  83
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/857,636
;      FILING DATE:  16-MAY-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/017,906
;      FILING DATE:  17-MAY-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PO0011
;      FILING DATE:  21-MAY-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  AU PO0363
;      FILING DATE:  07-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/019,765
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunter, Tom
;      REGISTRATION NUMBER:  38,498
;      REFERENCE/DOCKET NUMBER:  015280-278200US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  -
;      LOCATION:  1..22
;      OTHER INFORMATION:  /note= "PTCF21 primer"
US-08-857-636-20

Alignment Scores:
Pred. No.:                195             Length:         22
Score:                    6.00            Matches:        6
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              10.34%          Indels:         0
DB:                       4               Gaps:           0

US-10-031-158-14 (1-58) x US-08-857-636-20 (1-22)

Qy         28 ValPheLeuArgAsnPhe 33
              ||||||||||||||||||
Db         21 GTGTTCCTGAGAAATTTT 4

RESULT 4
US-09-396-196G-41485/c
; Sequence 41485, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
;  APPLICANT: Michael Mittmann
;  APPLICANT:  David Mack
;  APPLICANT:  David Lockhart
;  APPLICANT:  Affymetrix, Inc.
;  TITLE OF INVENTION: Methods of Genetic Analysis
;  FILE REFERENCE: 3101.1
;  CURRENT APPLICATION NUMBER: US/09/396,196G
;  CURRENT FILING DATE:  1999-09-15
```

; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41485
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-396-196G-41485

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-41485 (1-25)

```
Qy         46 ArgAlaThrArgPheTrp 51
              |||||||||||||||||
Db         24 CGTGCTACAAGGTTCTGG 7
```

RESULT 5
US-09-396-196G-41486/c
; Sequence 41486, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41486
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-396-196G-41486

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-41486 (1-25)

```
Qy         46 ArgAlaThrArgPheTrp 51
              |||||||||||||||||
Db         23 CGTGCTACAAGGTTCTGG 6
```

RESULT 6
US-09-396-196G-41487/c
; Sequence 41487, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41487
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-396-196G-41487

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-41487 (1-25)

```
Qy         46 ArgAlaThrArgPheTrp 51
              |||||||||||||||||
Db         22 CGTGCTACAAGGTTCTGG 5
```

RESULT 7
US-09-396-196G-41488/c
; Sequence 41488, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41488
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-396-196G-41488

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-41488 (1-25)

```
Qy         46 ArgAlaThrArgPheTrp 51
              |||||||||||||||||
Db         21 CGTGCTACAAGGTTCTGG 4
```

RESULT 8
US-09-396-196G-41489/c
; Sequence 41489, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1

; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41489
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-396-196G-41489

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-41489 (1-25)

```
Qy         46 ArgAlaThrArgPheTrp 51
              |||||||||||||||||||
Db         19 CGTGCTACAAGGTTCTGG 2
```

RESULT 9
US-09-396-196G-41490/c
; Sequence 41490, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 41490
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-396-196G-41490

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-41490 (1-25)

```
Qy         46 ArgAlaThrArgPheTrp 51
              |||||||||||||||||||
Db         18 CGTGCTACAAGGTTCTGG 1
```

RESULT 10
US-09-396-196G-113910/c
; Sequence 113910, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 113910
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-113910

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-113910 (1-25)

```
Qy         18 GlnSerSerArgArgLeu 23
              |||||||||||||||||||
Db         23 CAAAGCAGCCGTCGTCTT 6
```

RESULT 11
US-09-396-196G-127372
; Sequence 127372, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 127372
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-127372

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-127372 (1-25)

```
Qy         19 SerSerArgArgLeuGlu 24
              ||||||||||||||||||
Db          7 AGCAGTCGCAGACTCGAA 24
```

RESULT 12
US-09-396-196G-127373
; Sequence 127373, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.

; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 127373
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-127373

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-127373 (1-25)

```
Qy         19 SerSerArgArgLeuGlu 24
              ||||||||||||||||||
Db          4 AGCAGTCGCAGACTCGAA 21
```

RESULT 13
US-09-396-196G-127384
; Sequence 127384, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 127384
; LENGTH: 25
; TYPE: DNA
; ORGANISM: mus musculus
US-09-396-196G-127384

Alignment Scores:

Pred. No.:	220	Length:	25
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x US-09-396-196G-127384 (1-25)

```
Qy         19 SerSerArgArgLeuGlu 24
              ||||||||||||||||||
Db          1 AGCAGTCGCAGACTCGAA 18
```

RESULT 14
US-08-186-229-20/c
; Sequence 20, Application US/08186229
; Patent No. 5736316
; GENERAL INFORMATION:
; APPLICANT: Irvine, Bruce D.
; APPLICANT: Kolberg, Janice A.
; APPLICANT: Running, Joyce A.
; APPLICANT: Urdea, Michael S.
; TITLE OF INVENTION: HBV PROBES FOR USE IN SOLUTION
; TITLE OF INVENTION: PHASE SANDWICH HYBRIDIZATION ASSAYS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/186,229
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/813,586
; FILING DATE: 23-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Thomas E. Ciotti
; REGISTRATION NUMBER: 21,013
; REFERENCE/DOCKET NUMBER: 22300-20234.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-186-229-20

Alignment Scores:

Pred. No.:	261	Length:	30
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x US-08-186-229-20 (1-30)

```
Qy         18 GlnSerSerArgArgLeu 23
              ||||||||||||||||||
Db         21 CAGTCCTCGMGRAGATTG 4
```

RESULT 15
US-08-470-124-20/c
; Sequence 20, Application US/08470124
; Patent No. 5849481
; GENERAL INFORMATION:
; APPLICANT: Urdea, Michael S.
; APPLICANT: Horn, Thomas
; APPLICANT: Chang, Chu-An
; APPLICANT: Warner, Brian
; APPLICANT: Fultz, Timothy J.
; TITLE OF INVENTION: LARGE COMB-TYPE BRANCHED
; TITLE OF INVENTION: POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 87
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 545 Middlefield Road, Suite 200
; CITY: Menlo Park
; STATE: California
; COUNTRY: USA
; ZIP: 94025

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/470,124
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/813,588
;      FILING DATE:  23 December 1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ciotti, Thomas E.
;      REGISTRATION NUMBER:  21,013
;      REFERENCE/DOCKET NUMBER:  22300-20104.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-813-5600
;      TELEFAX:  415-327-2951
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-470-124-20

Alignment Scores:
Pred. No.:                261             Length:          30
Score:                    6.00            Matches:         6
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              10.34%          Indels:          0
DB:                       2               Gaps:            0

US-10-031-158-14 (1-58) x US-08-470-124-20 (1-30)

Qy          18 GlnSerSerArgArgLeu 23
               |||||||||||||||||||
Db          21 CAGTCCTCGMGRAGATTG 4


Search completed: October 16, 2005, 11:02:49
Job time : 131 secs
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 08:49:13 ; Search time 2981 Seconds
(without alignments)
740.600 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45551

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135843_19614/app_query.fasta_1.199
-DB=EST -QFMT=fastap -SUFFIX=p2n_olig_sz30.rst -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30
-USER=US10031158_@CGN_1_1_5180_@runat_14102005_135843_19614 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	6	10.3	28	8	AZ387839	AZ387839 1M0147H22
c	2	5	8.6	16	5	BQ583458	BQ583458 E011979-0
	3	5	8.6	19	8	AZ465954	AZ465954 1M0276E16
	4	5	8.6	20	2	AW246466	AW246466 2821777.3
	5	5	8.6	20	7	CF282165	CF282165 14ETL--09
	6	5	8.6	20	8	AZ832043	AZ832043 2M0112I11
	7	5	8.6	20	8	AZ957966	AZ957966 2M0225I01
	8	5	8.6	20	9	AG189044	AG189044 Pan trogl
c	9	5	8.6	20	9	CL660020	CL660020 PRI0135d_
	10	5	8.6	21	7	CF338234	CF338234 RCL1--01-
c	11	5	8.6	21	8	AZ331619	AZ331619 1M0059K09
c	12	5	8.6	21	8	AZ341108	AZ341108 1M0073O14
	13	5	8.6	21	8	AZ589400	AZ589400 1M0398C23
c	14	5	8.6	21	8	AZ806282	AZ806282 2M0068A15
c	15	5	8.6	21	8	AZ850337	AZ850337 2M0152H11
	16	5	8.6	21	8	AZ943299	AZ943299 2M0203K21
c	17	5	8.6	22	1	AI688330	AI688330 wc94c08.x
c	18	5	8.6	22	1	AJ661940	AJ661940 AJ661940
	19	5	8.6	22	7	CF310806	CF310806 ABF--05-K
	20	5	8.6	22	8	AZ317017	AZ317017 1M0035P09
c	21	5	8.6	22	9	TA231E08Q	AL480935 T. brucei
c	22	5	8.6	23	7	CF302431	CF302431 7LEAF--07
	23	5	8.6	23	8	AZ615086	AZ615086 1M0444C14
c	24	5	8.6	23	8	AZ819376	AZ819376 2M0089P23
c	25	5	8.6	23	9	TA173D05P	AL473896 T. brucei
c	26	5	8.6	24	1	AU257964	AU257964 AU257964
	27	5	8.6	24	7	CF295238	CF295238 30DGS--05
c	28	5	8.6	24	7	CF340367	CF340367 RCL1--07-
c	29	5	8.6	24	8	AZ852116	AZ852116 2M0154M12
c	30	5	8.6	24	9	TA155F12Q	AL472698 T. brucei
c	31	5	8.6	24	9	TA185C06P	AL474052 T. brucei
c	32	5	8.6	25	1	AI015571	AI015571 ov09f01.x
c	33	5	8.6	25	1	AI624473	AI624473 ts68d11.x
c	34	5	8.6	25	1	AJ663467	AJ663467 AJ663467
c	35	5	8.6	25	8	AZ997733	AZ997733 2M0284E22
	36	5	8.6	25	8	BH854637	BH854637 KG07068-5
c	37	5	8.6	25	9	TA232G09Q	AL481789 T. brucei
c	38	5	8.6	26	1	AJ666261	AJ666261 AJ666261
	39	5	8.6	26	1	AJ683031	AJ683031 AJ683031
c	40	5	8.6	26	1	AJ685374	AJ685374 AJ685374
	41	5	8.6	26	1	AU265518	AU265518 AU265518
c	42	5	8.6	26	8	AZ377143	AZ377143 1M0131O17
c	43	5	8.6	26	8	AZ447827	AZ447827 1M0245J08
	44	5	8.6	26	9	TA319G03Q	AL492820 T. brucei
	45	5	8.6	27	5	BQ592207	BQ592207 E012696-0

ALIGNMENTS

RESULT 1
AZ387839
LOCUS AZ387839 28 bp DNA linear GSS 02-OCT-2000
DEFINITION 1M0147H22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0147H22 R, genomic survey sequence.
ACCESSION AZ387839
VERSION AZ387839.1 GI:10501547
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 28)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0147 row: H column: 22
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 28.

```
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0147H22"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:              5.35e+03        Length:         28
Score:                  6.00            Matches:        6
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            10.34%          Indels:         0
DB:                     8               Gaps:           0

US-10-031-158-14 (1-58) x AZ387839 (1-28)

Qy          8 ProLeuPhePhePheLeu 13
              |||||||||||||||||
Db         11 CCATTATTTTTCTTTCTT 28

RESULT 2
BQ583458/c
LOCUS       BQ583458                 16 bp    mRNA    linear   EST 06-DEC-2002
DEFINITION  E011979-024-005-J11-SP6 MPIZ-ADIS-024-inflorescence Beta vulgaris
            cDNA clone 024-005-J11 5-PRIME, mRNA sequence.
ACCESSION   BQ583458
VERSION     BQ583458.1  GI:26113035
KEYWORDS    EST.
SOURCE      Beta vulgaris
  ORGANISM  Beta vulgaris
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllales; Amaranthaceae; Beta.
REFERENCE   1  (bases 1 to 16)
  AUTHORS   Herwig,R., Schulz,B., Weisshaar,B., Hennig,S., Steinfath,M.,
            Drungowski,M., Stahl,D., Wruck,W., Menze,A., O'Brien,J., Lehrach,H.
            and Radelof,U.
  TITLE     Construction of a 'unigene' cDNA clone set by oligonucleotide
            fingerprinting allows access to 25 000 potential sugar beet genes
  JOURNAL   Plant J. 32 (5), 845-857 (2002)
  MEDLINE   22362189
   PUBMED   12472698
COMMENT     Contact: Weisshaar B
            ADIS DNA core facility at MPIZ
            Max-Planck-Institute for Plant Breeding Research
            Carl-von-Linne Weg 10, 50829 Koeln, Germany
            Fax: 00492215062851
            Email: weisshaa@mpiz-koeln.mpg.de
            Insert Length: 16   Std Error: 0.00
            Plate: 5  row: J  column: 11
            Seq primer: SP6; CATACGATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .16
                     /organism="Beta vulgaris"
                     /mol_type="mRNA"
                     /cultivar="KWS2320 (double haploid, monogerm breeding
                     line)"
                     /db_xref="GABI:183240"
                     /db_xref="taxon:161934"
                     /clone="024-005-J11"
                     /tissue_type="inflorescence"
                     /lab_host="EMDH10B"
                     /clone_lib="MPIZ-ADIS-024-inflorescence"
                     /note="Vector: pCMVSPORT6; Site_1: SalI; Site_2: NotI;
                     cDNA library from sugar beet, library provided by KWS
                     Kleinwanzlebener Saatzucht AG Einbeck, Germany, contact:
                     b.schulz@kws.de; cloning sites SalI-NotI, primer sites and
                     orientation:
                     SP6-SalI-CCACGCGTCCG-5prime-cDNA-polyA-CC-NotI-T7;  Note:
                     Sequencing granted in the context of the GABI-Beet
                     project, local PI: Dr. Katharina Schneider, coordinator:
                     Prof. Christian Jung;  Sequence submission managed by
                     RZPD/GABI-Primary database: http://gabi.rzpd.de"
ORIGIN

Alignment Scores:
Pred. No.:              3.59e+04        Length:         16
Score:                  5.00            Matches:        5
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            8.62%           Indels:         0
DB:                     5               Gaps:           0

US-10-031-158-14 (1-58) x BQ583458 (1-16)

Qy          8 ProLeuPhePhePhe 12
              ||||||||||||||||
Db         15 CCCCTTTTTTTTTTT 1

RESULT 3
AZ465954
LOCUS       AZ465954                 19 bp    DNA     linear   GSS 04-OCT-2000
DEFINITION  1M0276E16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0276E16 F, genomic survey sequence.
ACCESSION   AZ465954
VERSION     AZ465954.1  GI:10624079
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0276  row: E  column: 16
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
```

High quality sequence stop: 19.

```
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0276E16"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN
```

Alignment Scores:

Pred. No.:	4.19e+04	Length:	19
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-14 (1-58) x AZ465954 (1-19)

```
Qy          8 ProLeuPhePhePhe 12
              |||||||||||||||
Db          2 CCCCTTTTTTTTTTT 16
```

```
RESULT 4
AW246466
LOCUS       AW246466                 20 bp    mRNA    linear   EST 07-JAN-2000
DEFINITION  2821777.3prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2821777 3',
            mRNA sequence.
ACCESSION   AW246466
VERSION     AW246466.1  GI:6589459
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: 2821777.5prime
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu Low Quality Sequence: 10
            contiguous PHRED high quality bases following vector sequence. Very
            Low Quality Sequence: Trace file contained 20 contiguous distinct
            peaks following vector sequence. Polyadenylation: Based upon the
            presence of a XhoI site followed by a run of 14 or more T residues
            at the beginning of the sequence, this cDNA insert was
            polyadenylated.
            Plate: LLCM7  row: L  column: 2
            High quality sequence stop: 10.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2821777"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_7"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
```

Alignment Scores:

Pred. No.:	4.38e+04	Length:	20
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x AW246466 (1-20)

```
Qy         11 PhePheLeuGlnLeu 15
              |||||||||||||||
Db          3 TTTTTTTTGCAGCTT 17
```

```
RESULT 5
CF282165
LOCUS       CF282165                 20 bp    mRNA    linear   EST 14-AUG-2003
DEFINITION  14ETL--09-I22.g1 Rice etiolated leaf plasmid cDNA library (14ETL)
            Oryza sativa (japonica cultivar-group) cDNA clone 14ETL--09-I22,
            mRNA sequence.
ACCESSION   CF282165
VERSION     CF282165.1  GI:33659552
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
  TITLE     Large-scale Sequencing Analysis of Rice ESTs
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, MyongJi University
            Yongin, Kyeonggi, Korea
            Tel: 82 31 330 6193
            Fax: 82 31 321 6355
            Email: bhnahm@ggbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES             Location/Qualifiers
     source          1. .20
```

```
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /db_xref="taxon:39947"
                     /clone="14ETL--09-I22"
                     /tissue_type="leaf"
                     /dev_stage="14 days after germination"
                     /lab_host="E.coli DH10B"
                     /clone_lib="Rice etiolated leaf plasmid cDNA library
                     (14ETL)"
                     /note="Vector: pCR4-TOPO; Site_1: EcoRI; mRNA was capped
                     with oligoribonucleotides and then used as templates for
                     RT-PCR."
ORIGIN

Alignment Scores:
Pred. No.:              4.38e+04        Length:         20
Score:                  5.00            Matches:        5
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            8.62%           Indels:         0
DB:                     7               Gaps:           0

US-10-031-158-14 (1-58) x CF282165 (1-20)

Qy          10 PhePhePheLeuGln 14
               |||||||||||||||
Db           4 TTTTTTTTTTTGCAA 18

RESULT 6
AZ832043
LOCUS       AZ832043                 20 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0112I11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0112I11 F, genomic survey sequence.
ACCESSION   AZ832043
VERSION     AZ832043.1  GI:13001951
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0112  row: I  column: 11
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 20.
FEATURES             Location/Qualifiers
     source          1...20
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0112I11"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:              4.38e+04        Length:         20
Score:                  5.00            Matches:        5
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            8.62%           Indels:         0
DB:                     8               Gaps:           0

US-10-031-158-14 (1-58) x AZ832043 (1-20)

Qy           9 LeuPhePhePheLeu 13
               |||||||||||||||
Db           2 CTATTTTTCTTCTTA 16

RESULT 7
AZ957966
LOCUS       AZ957966                 20 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0225I01F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0225I01 F, genomic survey sequence.
ACCESSION   AZ957966
VERSION     AZ957966.1  GI:13829193
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0225  row: I  column: 01
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 20.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
```

/clone="UUGC2M0225I01"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Alignment Scores:

Pred. No.:	4.38e+04	Length:	20
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-14 (1-58) x AZ957966 (1-20)

```
Qy          8 ProLeuPhePhePhe 12
              ||||||||||||||||
Db          6 CCCCTTTTTTTTTTT 20
```

RESULT 8
AG189044
LOCUS AG189044 20 bp DNA linear GSS 06-MAR-2004
DEFINITION Pan troglodytes DNA, clone: RP43-063F05.TJ, genomic survey sequence.
ACCESSION AG189044
VERSION AG189044.1 GI:45221220
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J., Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
TITLE BAC end sequences of Library RP-43
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 20)
AUTHORS Park,H., Kim,Y., Kim,S., Han,Y., Woo,T., Park,K., Eun,C.J., Hoon,S.T., Chu,M., Kim,H., Joo,S., Kim,C., Song,W. and Yoo,H.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-2002) Hong-Seog Park, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Genome Research Center (GRC); 52, Oun-dong, Yusong-gu, Daejeon 305-333, Korea (E-mail:redstone@mail.kribb.re.kr, URL:http://phs.grc.kribb.re.kr/, Tel:82-42-866-7181, Fax:82-42-860-4409)
COMMENT Clones are derived from the chimpanzee BAC library RP-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
FEATURES Location/Qualifiers
source 1. .20
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-063F05.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RP-43 Chimpanzee Male BAC Library"
ORIGIN

Alignment Scores:

Pred. No.:	4.38e+04	Length:	20
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	9	Gaps:	0

US-10-031-158-14 (1-58) x AG189044 (1-20)

```
Qy          8 ProLeuPhePhePhe 12
              ||||||||||||||||
Db          5 CCCCTTTTTTTTTTT 19
```

RESULT 9
CL660020/c
LOCUS CL660020 20 bp DNA linear GSS 09-JUL-2004
DEFINITION PRI0135d_H08 - PRI0135d.B21 (20) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.
ACCESSION CL660020
VERSION CL660020.1 GI:50144803
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 20)
AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppaDB: an AcedB database for the nematode satellite organism Pristionchus pacificus
JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
FEATURES Location/Qualifiers
source 1. .20
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus var. California"
/note="Vector: pEpifos-5 Fosmid vector"
ORIGIN

Alignment Scores:

Pred. No.:	4.38e+04	Length:	20
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	9	Gaps:	0

```
US-10-031-158-14 (1-58) x CL660020 (1-20)

Qy          5 ProProSerProLeu 9
              ||||||||||||||||
Db         16 CCCCCCTCCCCACTA 2

RESULT 10
CF338234
LOCUS       CF338234                 21 bp    mRNA    linear   EST 18-AUG-2003
DEFINITION  RCL1--01-A19.g1 Regenerated callus lambda phage cDNA library (RCL1)
            Oryza sativa (japonica cultivar-group) cDNA clone RCL1--01-A19,
            mRNA sequence.
ACCESSION   CF338234
VERSION     CF338234.1  GI:33824849
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
  TITLE     Large-scale Sequencing Analysis of Rice ESTs
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, MyongJi University
            Yongin, Kyeonggi, Korea
            Tel: 82 31 330 6193
            Fax: 82 31 321 6355
            Email: bhnahm@ggbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /db_xref="taxon:39947"
                     /clone="RCL1--01-A19"
                     /tissue_type="callus"
                     /dev_stage="proliferated callus on 2N6 media for 30 days"
                     /lab_host="E.coli SOLR"
                     /clone_lib="Regenerated callus lambda phage cDNA library
                     (RCL1)"
                     /note="Vector: pBluescript SK(+); Site_1: SstI; Site_2:
                     XhoI; cDNA was inserted into lamda Uni-ZAP XR vector at 5'
                     end with SstI and 3' end with XhoI site. Callus was
                     induced on 2N6 media for 30 days and cultured for 36hrs on
                     regenerated media"
ORIGIN

Alignment Scores:
Pred. No.:              4.58e+04        Length:         21
Score:                  5.00            Matches:        5
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            8.62%           Indels:         0
DB:                     7               Gaps:           0

US-10-031-158-14 (1-58) x CF338234 (1-21)

Qy          8 ProLeuPhePhePhe 12
              ||||||||||||||||
Db          5 CCCCTTTTTTTTTTT 19

RESULT 11
AZ331619/c
LOCUS       AZ331619                 21 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0059K09R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0059K09 R, genomic survey sequence.
ACCESSION   AZ331619
VERSION     AZ331619.1  GI:10394486
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0059  row: K  column: 09
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 21.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0059K09"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:              4.58e+04        Length:         21
Score:                  5.00            Matches:        5
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            8.62%           Indels:         0
DB:                     8               Gaps:           0

US-10-031-158-14 (1-58) x AZ331619 (1-21)

Qy          9 LeuPhePhePheLeu 13
              |||||||||||||||
Db         21 TTGTTTTTTTTTTTA 7

RESULT 12
AZ341108/c
```

```
LOCUS       AZ341108                 21 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0073O14F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0073O14 F, genomic survey sequence.
ACCESSION   AZ341108
VERSION     AZ341108.1  GI:10417030
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0073  row: O  column: 14
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 21.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0073O14"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:                4.58e+04        Length:          21
Score:                    5.00            Matches:         5
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              8.62%           Indels:          0
DB:                       8               Gaps:            0

US-10-031-158-14 (1-58) x AZ341108 (1-21)

Qy          8 ProLeuPhePhePhe 12
                  ||||||||||||||||
Db         19 CCCCTTTTTTTTTTT 5

RESULT 13
AZ589400
LOCUS       AZ589400                 21 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0398C23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0398C23 F, genomic survey sequence.
ACCESSION   AZ589400
VERSION     AZ589400.1  GI:11711590
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0398  row: C  column: 23
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 21.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0398C23"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:
Pred. No.:                4.58e+04        Length:          21
Score:                    5.00            Matches:         5
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              8.62%           Indels:          0
```

```
DB:                 8          Gaps:          0

US-10-031-158-14 (1-58) x AZ589400 (1-21)

Qy          8 ProLeuPhePhePhe 12
              ||||||||||||||||
Db          5 CCCCTTTTTTTTTTT 19

RESULT 14
AZ806282/c
LOCUS       AZ806282                 21 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0068A15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0068A15 F, genomic survey sequence.
ACCESSION   AZ806282
VERSION     AZ806282.1  GI:12967093
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0068  row: A  column: 15
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 21.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0068A15"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

Alignment Scores:

Pred. No.:              4.58e+04    Length:          21
Score:                  5.00        Matches:         5
Percent Similarity:     100.00%     Conservative:    0
Best Local Similarity:  100.00%     Mismatches:      0
Query Match:            8.62%       Indels:          0
DB:                     8           Gaps:            0

US-10-031-158-14 (1-58) x AZ806282 (1-21)

Qy          7 SerProLeuPhePhe 11
              |||||||||||||||
Db         16 TCCCCTCTTTTCTTT 2

RESULT 15
AZ850337/c
LOCUS       AZ850337                 21 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0152H11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0152H11 F, genomic survey sequence.
ACCESSION   AZ850337
VERSION     AZ850337.1  GI:13035243
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0152  row: H  column: 11
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 21.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0152H11"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
```

chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN

Alignment Scores:

Pred. No.:	4.58e+04	Length:	21
Score:	5.00	Matches:	5
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.62%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-14 (1-58) x AZ850337 (1-21)

```
Qy          6 ProSerProLeuPhe 10
              |||||||||||||||
Db         16 CCCTCCCCTCTCTTT 2
```

Search completed: October 16, 2005, 11:00:30
Job time : 2985 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 07:40:42 ; Search time 420 Seconds
(without alignments)
817.488 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3518947

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135842_19595/app_query.fasta_1.199
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=p2n_olig_sz30.rng -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30
-USER=US10031158_@CGN_1_1_708_@runat_14102005_135842_19595 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	8	13.8	27	4	AAF56408	Aaf56408 Human pro
c	2	8	13.8	27	4	AAF56409	Aaf56409 Human pro
c	3	7	12.1	21	2	AAQ04627	Aaq04627 CTG prime
	4	7	12.1	25	10	ADD28130	Add28130 HIV-1 LTR
c	5	6	10.3	19	6	ABX97507	Abx97507 Human NOV
c	6	6	10.3	19	12	ADN62410	Adn62410 Human NOV
	7	6	10.3	20	3	AAZ76949	Aaz76949 Human bia
	8	6	10.3	20	4	AAH23198	Aah23198 Human MMI
c	9	6	10.3	20	8	ABZ10328	Abz10328 Haematopo
c	10	6	10.3	20	10	ADB54415	Adb54415 PCR prime
c	11	6	10.3	20	10	ADC69900	Adc69900 Primer ol
	12	6	10.3	20	10	ADE77580	Ade77580 DRB3*0212
	13	6	10.3	20	10	ABZ85409	Abz85409 Human oli
	14	6	10.3	20	11	ABD21639	Abd21639 S100 calc
c	15	6	10.3	20	13	ADS89854	Ads89854 Human PCR
c	16	6	10.3	22	2	AAV15960	Aav15960 NBCCS (PT
c	17	6	10.3	23	6	AAD39184	Aad39184 Mouse 572
	18	6	10.3	23	10	ADD28129	Add28129 HIV-1 LTR
	19	6	10.3	25	9	ACI78562	Aci78562 Human mic
c	20	6	10.3	25	9	ACK21141	Ack21141 Human mic
	21	6	10.3	25	9	ACI31211	Aci31211 Human mic
c	22	6	10.3	25	9	ACI67768	Aci67768 Human mic
c	23	6	10.3	25	9	ACI83555	Aci83555 Human mic
	24	6	10.3	25	9	ACI78563	Aci78563 Human mic
	25	6	10.3	25	10	ADB88664	Adb88664 Frizzled-
	26	6	10.3	26	8	ABX94983	Abx94983 Barley RA
	27	6	10.3	26	8	ABX94962	Abx94962 Barley Ra
	28	6	10.3	27	12	ADP70068	Adp70068 Novel mam
c	29	6	10.3	28	10	ADF43026	Adf43026 Bacterial
c	30	6	10.3	30	2	AAQ45800	Aaq45800 HBV ampli
c	31	6	10.3	30	2	AAV07797	Aav07797 HBV.71 am
c	32	6	10.3	30	2	AAV83026	Aav83026 Amplifier
c	33	6	10.3	30	6	ABX69492	Abx69492 Novel Hel
	34	6	10.3	30	10	ADC10250	Adc10250 Human NOV
c	35	6	10.3	30	12	ADM18572	Adm18572 Human sub
	36	6	10.3	30	12	ADP08610	Adp08610 PCR prime
c	37	5	8.6	15	4	AAH46713	Aah46713 Type 11 p
c	38	5	8.6	15	4	ABA02603	Aba02603 PTEN targ
	39	5	8.6	15	4	AAF46222	Aaf46222 IGFBP2 ol
c	40	5	8.6	15	4	AAF52874	Aaf52874 IGF-I oli
c	41	5	8.6	16	2	AAQ25455	Aaq25455 Purine ri
c	42	5	8.6	16	2	AAQ44151	Aaq44151 Sequence
c	43	5	8.6	16	2	AAQ43987	Aaq43987 HIV-1 tri
c	44	5	8.6	16	2	AAQ45376	Aaq45376 Polypurin
c	45	5	8.6	16	2	AAQ95127	Aaq95127 Oligonucl

ALIGNMENTS

RESULT 1
AAF56408/c
ID AAF56408 standard; DNA; 27 BP.
XX
AC AAF56408;
XX
DT 12-APR-2001 (first entry)
XX
DE Human prostate specific TCRgamma transcript PCR primer TCRCgamma-R2.
XX
KW Human; TARP; prostate cancer; breast cancer; immunotherapy;
KW T cell receptor gamma alternate reading frame protein; TCRgamma;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200104309-A1.
XX
PD 18-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US019039.
XX
PR 13-JUL-1999; 99US-0143560P.
PR 01-OCT-1999; 99US-0157471P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Pastan I, Essand M, Lee B, Vasmatzis G, Wolfgang C;
XX

DR WPI; 2001-081050/09.
XX
PT Isolated T-cell receptor gamma alternate reading frame protein useful for
PT diagnosing and raising an immune response to prostate cancer and breast
PT cancer.
XX
PS Disclosure; Fig 8; 85pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human T cell receptor alternate reading frame protein (TARP). This
CC protein is expressed in prostate and breast cancer cells at higher levels
CC than normal and so can be used in the immunotherapy of these cancers, as
CC well as their detection and prevention
XX
SQ Sequence 27 BP; 6 A; 4 C; 8 G; 9 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	12	Length:	27
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	13.79%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x AAF56408 (1-27)

```
Qy         13 LeuGlnLeuLeuLysGlnSerSer 20
              |||||||||||||||||||||||||
Db         26 CTTCAATTGCTGAAACAAAGCTCC 3
```

RESULT 2
AAF56409/c
ID AAF56409 standard; DNA; 27 BP.
XX
AC AAF56409;
XX
DT 12-APR-2001 (first entry)
XX
DE Human prostate specific TCRgamma transcript PCR primer TCRCgamma-R3.
XX
KW Human; TARP; prostate cancer; breast cancer; immunotherapy;
KW T cell receptor gamma alternate reading frame protein; TCRgamma;
KW PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200104309-A1.
XX
PD 18-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US019039.
XX
PR 13-JUL-1999; 99US-0143560P.
PR 01-OCT-1999; 99US-0157471P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Pastan I, Essand M, Lee B, Vasmatzis G, Wolfgang C;
XX
DR WPI; 2001-081050/09.
XX
PT Isolated T-cell receptor gamma alternate reading frame protein useful for
PT diagnosing and raising an immune response to prostate cancer and breast
PT cancer.
XX
PS Disclosure; Fig 8; 85pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human T cell receptor alternate reading frame protein (TARP). This
CC protein is expressed in prostate and breast cancer cells at higher levels
CC than normal and so can be used in the immunotherapy of these cancers, as
CC well as their detection and prevention
XX
SQ Sequence 27 BP; 10 A; 6 C; 6 G; 5 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	12	Length:	27
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	13.79%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x AAF56409 (1-27)

```
Qy         23 LeuGluHisThrPheValPheLeu 30
              |||||||||||||||||||||||||
Db         26 CTGGAACATACCTTTGTCTTCTTG 3
```

RESULT 3
AAQ04627/c
ID AAQ04627 standard; DNA; 21 BP.
XX
AC AAQ04627;
XX
DT 25-MAR-2003 (revised)
DT 31-OCT-2002 (revised)
DT 02-OCT-1990 (first entry)
XX
DE CTG primer for T-cell receptor gene.
XX
KW Primer; T-cell receptor; C-region; gamma chain; lymphoid leukaemia;
KW lymphoma; PCR; ss.
XX
OS Synthetic.
XX
PN WO9004648-A.
XX
PD 03-MAY-1990.
XX
PF 20-OCT-1988; 88AU-00001057.
XX
PR 20-OCT-1988; 88AU-00001057.
XX
PA (MORL/) MORLEY A A.
XX
PI Brisco MJ;
XX
DR WPI; 1990-164030/21.
XX
PT Detection of lymphoid leukaemia and/or lymphoma - by determn. of
PT homogeneity or heterogeneity of the length of immunoglobulin and/or T-
PT receptor gene segments.
XX
PS Claim 12; Page 18; 23pp; English.
XX
CC The primer binds to the C-region of T-receptor gamma chain gene. Using
CC the primer T-receptor gene segments can be amplified (by PCR). Size sepn.
CC of the amplified fragments can then be used to determine homogeneity or
CC heterogeneity. This provides a rapid and sensitive diagnostic test for
CC leukaemia/lymphoma by detection of malignancy and determn.of the B- or T-
CC lymphocyte origin of the tumour. See also AAQ04603-Q04640. (Updated on 31
CC -OCT-2002 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
SQ Sequence 21 BP; 4 A; 4 C; 5 G; 8 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	86.1	Length:	21
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.07%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x AAQ04627 (1-21)

```
Qy          15 LeuLeuLysGlnSerSerArg 21
               |||||||||||||||||||||
Db          21 TTGCTGAAACAAAGCTCCAGA 1

RESULT 4
ADD28130
ID   ADD28130 standard; DNA; 25 BP.
XX
AC   ADD28130;
XX
DT   15-JAN-2004  (first entry)
XX
DE   HIV-1 LTR region DNA fragment PNAIa.
XX
KW   conjugate; peptide nucleic acid; PNA; transport mediator;
KW   addressing protein; nuclear importation; penetratin family; tat gene;
KW   rev gene; anti-HIV; primer; ss.
XX
OS   Human immunodeficiency virus 1.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    19. .21
FT                   /*tag=  a
FT                   /note= "DNA sequence interrupted by an undisclosed
FT                   peptide linker"
FT   modified_base   25
FT                   /*tag=  b
FT                   /mod_base= OTHER
FT                   /note= "optionally pseudo-isocytosine or cytosine"
XX
PN   WO2003006065-A2.
XX
PD   23-JAN-2003.
XX
PF   12-JUL-2002; 2002WO-DE002564.
XX
PR   12-JUL-2001; 2001DE-01033307.
XX
PA   (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM STIFTUNG.
XX
PI   Braun K,  Waldeck W,  Pipkorn R,  Braun I,  Debus J;
XX
DR   WPI; 2003-256334/25.
XX
PT   New conjugate for delivery to the cell nucleus, useful for treating
PT   infection by human immune deficiency virus, comprises targeting peptides
PT   and peptide nucleic acid.
XX
PS   Disclosure; Page 8; 30pp; German.
XX
CC   This invention describes a novel conjugate for mediating specific
CC   transport to the cell nucleus of a peptide nucleic acid (PNA) that
CC   hybridizes to an HIV (human immune deficiency virus) gene, or part of it.
CC   The conjugate comprises a transport mediator for the cell membrane, an
CC   addressing protein or peptide for nuclear importation and the PNA being
CC   delivered. The transport mediator is a peptide or protein able to pass
CC   through the plasma membrane, particularly derived from the penetratin
CC   family protein. The PNA is designed to hybridize to at least part of the
CC   tat or rev genes and is especially targeted to the long terminal repeat
CC   or the polypurine tract, the central DNA flap, Nef, NCp7 or the gag slice
CC   site, at the RNA level. The conjugate has the structure transport
CC   mediator-addressing protein-PNA, particularly with a spacer (e.g.
CC   polylysine or a synthetic polymer) between addressing protein and PNA.
CC   The conjugate of the invention has anti-HIV activity. The transport
CC   mediator/addressing protein combination provide rapid and efficient
CC   transport of PNA to the nucleus and PNA is resistant to proteases and
CC   nucleases. The conjugate can be used in both early and late stages of
CC   infection. This sequence represents a fragment of HIV-1 LTR region used
CC   to construct PNA's used in the method of the invention.
XX
SQ   Sequence 25 BP; 2 A; 9 C; 0 G; 14 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                101          Length:          25
Score:                    7.00         Matches:         7
Percent Similarity:       100.00%      Conservative:    0
Best Local Similarity:    100.00%      Mismatches:      0
Query Match:              12.07%       Indels:          0
DB:                       10           Gaps:            0

US-10-031-158-14 (1-58) x ADD28130 (1-25)

Qy           7 SerProLeuPhePhePheLeu 13
                |||||||||||||||||||||
Db           4 TCCCCCCTTTTCTTTTTTCTT 24

RESULT 5
ABX97507/c
ID   ABX97507 standard; DNA; 19 BP.
XX
AC   ABX97507;
XX
DT   20-MAY-2003  (first entry)
XX
DE   Human NOV-associated forward primer from primer-probe set Ag3505.
XX
KW   NOVX; cytostatic; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
KW   hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
KW   human; PCR; primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200272757-A2.
XX
PD   19-SEP-2002.
XX
PF   08-MAR-2002; 2002WO-US006908.
XX
PR   08-MAR-2001; 2001US-0274101P.
PR   08-MAR-2001; 2001US-0274194P.
PR   08-MAR-2001; 2001US-0274281P.
PR   08-MAR-2001; 2001US-0274322P.
PR   09-MAR-2001; 2001US-0274849P.
PR   12-MAR-2001; 2001US-0275235P.
PR   13-MAR-2001; 2001US-0275578P.
PR   13-MAR-2001; 2001US-0275579P.
PR   13-MAR-2001; 2001US-0275601P.
PR   14-MAR-2001; 2001US-0276000P.
PR   16-MAR-2001; 2001US-0276776P.
PR   19-MAR-2001; 2001US-0276994P.
PR   20-MAR-2001; 2001US-0277239P.
PR   20-MAR-2001; 2001US-0277321P.
PR   20-MAR-2001; 2001US-0277327P.
PR   21-MAR-2001; 2001US-0277791P.
PR   22-MAR-2001; 2001US-0277833P.
PR   23-MAR-2001; 2001US-0278152P.
PR   26-MAR-2001; 2001US-0278894P.
PR   27-MAR-2001; 2001US-0278999P.
PR   27-MAR-2001; 2001US-0279036P.
PR   28-MAR-2001; 2001US-0279344P.
PR   30-MAR-2001; 2001US-0277338P.
PR   30-MAR-2001; 2001US-0279995P.
PR   30-MAR-2001; 2001US-0280233P.
PR   02-APR-2001; 2001US-0280802P.
PR   02-APR-2001; 2001US-0280822P.
PR   02-APR-2001; 2001US-0280900P.
PR   04-APR-2001; 2001US-0281194P.
PR   13-APR-2001; 2001US-0283675P.
PR   30-APR-2001; 2001US-0287424P.
PR   02-MAY-2001; 2001US-0288066P.
PR   03-MAY-2001; 2001US-0288342P.
PR   03-MAY-2001; 2001US-0288528P.
PR   15-MAY-2001; 2001US-0291190P.
PR   16-MAY-2001; 2001US-0291099P.
PR   16-MAY-2001; 2001US-0291240P.
```

PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 07-MAR-2002; 2002US-00092900.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y, Anderson D;
PI Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H, Alsobrook JP;
PI Lepley DM, Rieger DK;
XX
DR WPI; 2002-723332/78.
XX
PT NOVX polypeptides and polynucleotides, useful for preventing or treating
PT a disorder associated with aberrant NOVX expression or activity e.g.,
PT cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
PT asthma.
XX
PS Example C; Page 957; 1103pp; English.
XX
CC This invention describes novel human NOVX polypeptides which have
CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and hypotensive
CC activity. Pharmaceutical compositions comprising the NOVX proteins or
CC nucleic acid molecules or NOVX antibodies are useful for preventing or
CC treating a disorder associated with aberrant NOVX expression or activity
CC e.g. cancer, hypertension, atherosclerosis, cardiomyopathy or bronchial
CC asthma. The products of the invention can be used for gene therapy or in
CC a vaccine. ABX13460-ABX13462 and ABX97186-ABX97593 represent PCR primers
CC and probes used in the amplification and isolation of the NOVX
CC polynucleotides represented in ABX97008-ABX97185 which encode the
CC polypeptides represented in ABU65041-ABU65218
XX
SQ Sequence 19 BP; 4 A; 6 C; 5 G; 4 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	714	Length:	19
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	6	Gaps:	0

US-10-031-158-14 (1-58) x ABX97507 (1-19)

```
Qy         48 ThrArgPheTrpAspPro 53
              |||||||||||||||||
Db         19 ACTAGATTCTGGGACCCG 2
```

RESULT 6
ADN62410/c
ID ADN62410 standard; DNA; 19 BP.
XX
AC ADN62410;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human NOV98a RTQ-PCR forward primer.
XX
KW Human; ss; PCR; NOVX; diabetes; obesity; infectious disease; anorexia;
KW cancer-associated cachexia; cancer; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; chronic disease; primer; RTQ-PCR;
KW real time quantitative PCR.
XX
OS Homo sapiens.
XX
PN US2004043382-A1.
XX
PD 04-MAR-2004.
XX
PF 07-MAR-2002; 2002US-00092900.
XX
PR 08-MAR-2001; 2001US-0274191P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281444P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.

PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
XX
PA (PADI/) PADIGARU M.
PA (SPYT/) SPYTEK K A.
PA (SHEN/) SHENOY S G.
PA (TAUP/) TAUPIER R J.
PA (PENA/) PENA C E A.
PA (LILL/) LI L.
PA (ZERH/) ZERHUSEN B D.
PA (GUSE/) GUSEV V Y.
PA (JIWW/) JI W.
PA (GORM/) GORMAN L.
PA (MILL/) MILLER C E.
PA (KEKU/) KEKUDA R.
PA (PATT/) PATTURAJAN M.
PA (GANG/) GANGOLLI E A.
PA (VERN/) VERNET C A M.
PA (GUOX/) GUO X S.
PA (TCHE/) TCHERNEV V T.
PA (FERN/) FERNANDES E R.
PA (CASM/) CASMAN S J.
PA (MALY/) MALYANKAR U M.
PA (GERL/) GERLACH V.
PA (LIUY/) LIU Y.
PA (ANDE/) ANDERSON D W.
PA (SPAD/) SPADERNA S K.
PA (CATT/) CATTERTON E.
PA (LEIT/) LEITE M W.
PA (ZHON/) ZHONG H.
PA (ALSO/) ALSOBROOK J P.
PA (LEPL/) LEPLEY D M.
PA (RIEG/) RIEGER D K.
PA (BURG/) BURGESS C E.
XX
PI Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L;
PI Zerhusen BD, Gusev VY, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangolli EA, Vernet CAM, Guo XS, Tchernev VT;
PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
PI Anderson DW, Spaderna SK, Catterton E, Leite MW, Zhong H;
PI Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2004-225693/21.
XX
PT New NOVX polypeptides and nucleic acid molecules useful for diagnosing,
PT preventing or treating NOVX-associated disorders, e.g. cancer, diabetes,
PT infection or obesity, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Example C; SEQ ID NO 680; 786pp; English.
XX
CC The invention relates to an isolated polypeptide (designated NOVX, or
CC NOV1-NOV127) comprising a sequence selected from 178 fully defined amino
CC acid sequences (and their mature forms, variants and fragments). Also
CC included are an isolated nucleic acid molecule encoding NOVx, a vector
CC comprising the nucleic acid, a cell comprising the vector, methods for
CC determining the presence or amount of the polypeptide or the nucleic acid
CC molecule in a sample, methods for determining the presence of or
CC predisposition to a disease associated with altered levels of expression
CC of the above polypeptide or nucleic acid molecule in a first mammalian
CC subject, a method for identifying an agent that binds to the above
CC polypeptide, a method for identifying a potential therapeutic agent for
CC use in the treatment of a pathology that is related to aberrant
CC expression or physiological interactions of the polypeptide, a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide and a method for modulating
CC the activity of the polypeptide cited above. The composition and methods
CC are useful for diagnosing, preventing or treating diseases such as
CC diabetes, obesity, infectious diseases, anorexia, cancer-associated
CC cachexia, cancer, neurodegenerative disorders like Alzheimer's disease or
CC Parkinson's disease, immune disorders, haematopoietic disorders,
CC dyslipidaemias, and other chronic diseases. These may also be used in
CC chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The polypeptides are also useful as vaccines. The
CC present sequence is an RTQ-PCR (real time quantitative PCR) primer used
CC to assay tissue specific expression of a NOVX mRNA.
XX
SQ Sequence 19 BP; 4 A; 6 C; 5 G; 4 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	714	Length:	19
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	12	Gaps:	0

US-10-031-158-14 (1-58) x ADN62410 (1-19)

```
Qy      48 ThrArgPheTrpAspPro 53
           |||||||||||||||||
Db      19 ACTAGATTCTGGGACCCG 2
```

RESULT 7
AAZ76949
ID AAZ76949 standard; DNA; 20 BP.
XX
AC AAZ76949;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human biallelic marker downstream amplification primer SEQ ID NO:11305.
XX
KW Human genome; biallelic marker; high density disequilibrium map;
KW genomic map; haplotype; phenotype; polymorphic base; genotyping;
KW haplotyping; hybridisation; identification; characterisation;
KW amplification; single nucleotide polymorphism; SNP; PCR primer;
KW diagnosis; ss.
XX
OS Homo sapiens.
XX
PN WO9954500-A2.
XX
PD 28-OCT-1999.
XX
PF 21-APR-1999; 99WO-IB000822.
XX
PR 21-APR-1998; 98US-0082614P.
PR 23-NOV-1998; 98US-0109732P.
XX
PA (GEST) GENSET.
XX
PI Cohen D, Blumenfeld M, Chumakov I;
XX
DR WPI; 2000-013267/01.
XX
PT Novel biallelic markers used to construct a high density disequilibrium
PT map of the human genome.
XX
PS Claim 9; Page 2640; 2745pp; English.
XX
CC AAZ65654 to AAZ69578 represent human biallelic markers from the present
CC invention, which contain a polymorphic base at position 24 of their
CC nucleotide sequences. AAZ69579 to AAZ77440 represent amplification
CC primers for the biallelic markers. The biallelic markers of the invention

CC have a variety of uses: they can be used for high density mapping of the
CC human genome, and in complex association studies and haplotyping studies
CC which are useful in determining the genetic basis for disease states.
CC Compositions and methods of the invention can also be useful for the
CC identification of the targets for the development of pharmaceutical
CC agents and diagnostic methods, as well as the characterisation of the
CC differential efficacious responses to and side effects from
CC pharmaceutical agents acting on a disease as well as other treatment.
CC N.B. The SEQ ID NOs 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and
CC 3367, are not actually given a sequence in the Sequence Listing from the
CC present invention
XX
SQ Sequence 20 BP; 2 A; 5 C; 2 G; 11 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	3	Gaps:	0

US-10-031-158-14 (1-58) x AAZ76949 (1-20)

```
Qy          8 ProLeuPhePhePheLeu 13
              |||||||||||||||||
Db          3 CCCCTATTTTTTTTCTG 20
```

RESULT 8
AAH23198
ID AAH23198 standard; DNA; 20 BP.
XX
AC AAH23198;
XX
DT 17-SEP-2001 (first entry)
XX
DE Human MMIF mRNA inhibiting antisense oligo ISIS #112694.
XX
KW Macrophage migration inhibitory factor; MMIF; antisense; neurological;
KW hyperproliferation; nootropic; antihormonal; immunosuppresive; human;
KW antiinflammatory; cytostatic; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO200153317-A1.
XX
PD 26-JUL-2001.
XX
PF 16-JAN-2001; 2001WO-US001475.
XX
PR 20-JAN-2000; 2000US-00489869.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Murray SF, Cowsert LM, Wyatt JR;
XX
DR WPI; 2001-451899/48.
XX
PT New antisense compound(s) are useful to inhibit a nucleic acid molecule
PT encoding macrophage migration inhibitory factor.
XX
PS Claim 3; Page 82; 105pp; English.
XX
CC The invention relates to antisense oligonucleotides 8-30 nucleotides in
CC length targeted to a nucleic acid molecule encoding macrophage migration
CC inhibitory factor (MMIF), where the antisense compound specifically
CC hybridizes with and inhibits the expression of MMIF. The antisense
CC nucleotides are useful for the treatment of a disease or condition
CC associated with MMIF such as neurological, hormonal, immune, inflammatory
CC or hyperproliferative disorder. Sequences AAH23191-268 represent chimeric
CC antisense phosphorothioate oligonucleotides used for inhibition of human
CC MMIF mRNA expression
XX
SQ Sequence 20 BP; 6 A; 7 C; 7 G; 0 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x AAH23198 (1-20)

```
Qy         52 AspProArgArgGlyThr 57
              ||||||||||||||||||
Db          3 GACCCGCGCAGAGGCACA 20
```

RESULT 9
ABZ10328/c
ID ABZ10328 standard; DNA; 20 BP.
XX
AC ABZ10328;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related primer SEQ ID NO:468.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; probe; primer; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 11; Page 36; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA

CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 20 BP; 8 A; 0 C; 11 G; 1 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	8	Gaps:	0

US-10-031-158-14 (1-58) x ABZ10328 (1-20)

```
Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1
```

RESULT 10
ADB54415/c
ID ADB54415 standard; DNA; 20 BP.
XX
AC ADB54415;
XX
DT 04-DEC-2003 (first entry)
XX
DE PCR primer 83 used to amplify genomic DNA region.
XX
KW colon cell proliferative disorder; non methylated CpG dinucleotide;
KW cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ss;
KW PCR; primer.
XX
OS Unidentified.
XX
PN WO2003072821-A2.
XX
PD 04-SEP-2003.
XX
PF 27-FEB-2003; 2003WO-EP002035.
XX
PR 27-FEB-2002; 2002EP-00004551.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Adorjan P, Burger M, Maier S, Nimmrich I, Becker E, Lesche R;
PI Rujan T, Schmitt A;
XX
DR WPI; 2003-731620/69.
XX
PT Detecting and differentiating between colon cell proliferative disorders
PT associated with a gene or its regulatory regions comprises contacting a
PT target nucleic acid in a biological sample obtained from the subject with
PT a reagent.
XX
PS Claim 15; Page 24; 74pp; English.
XX
CC The invention relates to a novel method for detecting and differentiating
CC between colon cell proliferative disorders associated with at least one
CC gene or its regulatory regions. The method comprises contacting a target
CC nucleic acid in a biological sample obtained from the subject with at
CC least one reagent or a series of reagents, where the reagent or series of
CC reagents, distinguishes between methylated and non methylated CpG
CC dinucleotides within the target nucleic acid. The molecules of the
CC invention demonstrate cytostatic activity whilst the method may useful
CC for detecting and differentiating between colon cell proliferative
CC disorders, including cancers such as colon adenoma and colon carcinoma.
CC The PNA (peptide nucleic acid)-oligomers are useful as probes for
CC determining cytosine methylation state or single nucleotide
CC polymorphisms. The current sequence is that of the PCR primer of the
CC invention which was used to amplify the genomic DNA region.
XX
SQ Sequence 20 BP; 8 A; 0 C; 11 G; 1 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	10	Gaps:	0

US-10-031-158-14 (1-58) x ADB54415 (1-20)

```
Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1
```

RESULT 11
ADC69900/c
ID ADC69900 standard; DNA; 20 BP.
XX
AC ADC69900;
XX
DT 18-DEC-2003 (first entry)
XX
DE Primer oligo used to amplify pretreated genomic DNA (SeqID 389).
XX
KW PCR; primer; ss; lung cell proliferative disorder; CpG dinucleotide;
KW adenocarcinoma; squamous cell carcinoma; cytostatic; probe; PNA-oligomer;
KW cytosine methylation state.
XX
OS Unidentified.
XX
PN WO2003052135-A2.
XX
PD 26-JUN-2003.
XX
PF 10-DEC-2002; 2002WO-EP014026.
XX
PR 14-DEC-2001; 2001DE-01061625.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Burger M, Field JK, Genc B, Liloglou T, Lipscher E, Maier S;
PI Nimmrich I;
XX
DR WPI; 2003-533029/50.
XX
PT Detecting and differentiating cytosine methylation state of genomic DNA,
PT useful for diagnosing, treating prognosticating and/or monitoring lung
PT cell proliferative disorders e.g. adenocarcinoma and squamous cell
PT carcinoma.
XX
PS Claim 11; SEQ ID NO 389; 58pp; English.
XX
CC This invention relates to a novel method for detecting and
CC differentiating between lung cell proliferative disorders associated with
CC at least one gene and/or their regulatory regions. Specifically, it
CC refers to a method comprising contacting a target nucleic acid in a
CC biological sample with at least one reagent, wherein the reagent is able
CC to distinguish between methylated and non-methylated CpG dinucleotides
CC present in the target DNA. As such, it is possible to further
CC differentiate and diagnose medical conditions including adenocarcinoma
CC and squamous cell carcinoma, and their respective adjacent lung tissue.
CC The present invention describes cytostatic oligomers and PNA-oligomers
CC that are useful as probes for determining the cytosine methylation state
CC or single nucleotide polymorphisms (SNPs) of the target sequence. This
CC oligonucleotide sequence is a primer oligomer used for the amplification
CC of pretreated DNA (i.e. where unmethylated cytosine bases are converted
CC to uracil), used in an exemplification of the invention.
XX

```
SQ   Sequence 20 BP; 8 A; 0 C; 11 G; 1 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              749          Length:         20
Score:                  6.00         Matches:        6
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            10.34%       Indels:         0
DB:                     10           Gaps:           0

US-10-031-158-14 (1-58) x ADC69900 (1-20)

Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1

RESULT 12
ADE77580
ID   ADE77580 standard; DNA; 20 BP.
XX
AC   ADE77580;
XX
DT   29-JAN-2004  (first entry)
XX
DE   DRB3*0212 probe designed to analyse the HLA-DRB polymorphic region.
XX
KW   HLA-DRB; probe; ss; human; multiplexed elongation assay;
KW   enzymatic recognition;
KW   cystic fibrosis conductance transmembrane regulator; CFTR;
KW   human leukocyte antigen; HLA; genetic testing; carrier screening;
KW   genotyping; profiling; polymorphic.
XX
OS   Homo sapiens.
XX
PN   WO2003034029-A2.
XX
PD   24-APR-2003.
XX
PF   15-OCT-2002; 2002WO-US033012.
XX
PR   15-OCT-2001; 2001US-0329427P.
PR   15-OCT-2001; 2001US-0329428P.
PR   15-OCT-2001; 2001US-0329619P.
PR   15-OCT-2001; 2001US-0329620P.
PR   14-MAR-2002; 2002US-0364416P.
XX
PA   (BIOA-) BIOARRAY SOLUTIONS LTD.
XX
PI   Li AX,  Hashmi G,  Seul M;
XX
DR   WPI; 2003-393553/37.
XX
PT   Concurrent interrogation of a number of polymorphic sites, useful for
PT   genetic testing, carrier screening, genetic profiling, and identity
PT   testing, comprises conducting a multiplexed elongation assay using
PT   probes.
XX
PS   Example 2; Page 38; 143pp; English.
XX
CC   This invention relates to a novel method for the concurrent interrogation
CC   of a number of polymorphic sites in the presence of, and without
CC   interference from, non-designated polymorphic sites. Specifically, it
CC   comprises conducting a multiplexed elongation assay by applying one or
CC   more temperature cycles to achieve linear amplification of the target or
CC   a combination of annealing and elongation steps under temperature-
CC   controlled conditions. Furthermore, this detection method uses probe
CC   extension or elongation and relies on enzymatic recognition, a superior
CC   technique that no longer depends on differential hybridisation. The
CC   present invention describes probes and methods useful for identifying or
CC   detecting polymorphisms at one or more designated sites, such that they
CC   can identify mutations within the cystic fibrosis conductance
CC   transmembrane regulator (CFTR) or the human leukocyte antigen (HLA)
CC   genes. In addition, concurrent interrogation of a multiplicity of
CC   polymorphic sites is useful for genetic testing, carrier screening,
CC   genotyping or genetic profiling, and identity testing. This
CC   oligonucleotide is a probe designed to analyse the polymorphic region of
CC   the HLA-DRB gene of the invention.
XX
SQ   Sequence 20 BP; 3 A; 5 C; 4 G; 8 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              749          Length:         20
Score:                  6.00         Matches:        6
Percent Similarity:     100.00%      Conservative:   0
Best Local Similarity:  100.00%      Mismatches:     0
Query Match:            10.34%       Indels:         0
DB:                     10           Gaps:           0

US-10-031-158-14 (1-58) x ADE77580 (1-20)

Qy         12 PheLeuGlnLeuLeuLys 17
              ||||||||||||||||||
Db          1 TTCTTGCAGCTGCTTAAG 18

RESULT 13
ABZ85409
ID   ABZ85409 standard; DNA; 20 BP.
XX
AC   ABZ85409;
XX
DT   17-OCT-2003  (first entry)
XX
DE   Human oligonucleotide sequence.
XX
KW   Human; antisense; lung dysfunction; nasal airway dysfunction;
KW   antiinflammatory steroid; ubiquinone; antiinflammatory; antiallergic;
KW   antiasthmatic; hypotensive; immunosuppressive; cytostatic; gene therapy;
KW   antisense gene therapy; respiratory; lung; adenosine sensitivity;
KW   adenosine receptor; bronchodilation; bronchoconstriction; lung allergy;
KW   lung inflammation; respiratory disease; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285308-A2.
XX
PD   31-OCT-2002.
XX
PF   23-APR-2002; 2002WO-US013135.
XX
PR   24-APR-2001; 2001US-0286137P.
XX
PA   (EPIG-) EPIGENESIS PHARM INC.
XX
PI   Nyce JW,  Li Y,  Sandrasagra A,  Katz E,  Pabalan J,  Aguilar D;
PI   Miller S,  Tang L,  Shahabuddin S;
XX
DR   WPI; 2003-229219/22.
XX
PT   Pharmaceutical composition for treating ailments associated with impaired
PT   respiration, has oligo(s) antisense to specific gene(s) or its
PT   corresponding RNAs, and glucocorticoid or non-glucocorticoid steroid or
PT   ubiquinone.
XX
PS   Claim 15; SEQ ID NO 651; 872pp; English.
XX
CC   The invention relates to a novel pharmaceutical composition, which has a
CC   first active agent comprising an oligonucleotide antisense to the
CC   initiation codon, coding region, 5' or 3' end genomic flanking regions,
CC   5' and 3' intron-exon junctions, or regions within 2-10 nucleotides of
CC   junctions of genes encoding a polypeptide associated with lung and/or
CC   nasal airway dysfunction and a second active agent comprising an
CC   antiinflammatory steroid and ubiquinone. A composition of the invention
CC   has antiinflammatory, antiallergic, antiasthmatic, hypotensive,
CC   immunosuppressive, and cytostatic activity. The composition may have a
CC   use in antisense gene therapy. The composition is useful for treating or
CC   preventing a respiratory, lung or malignant disease or condition, also
```

CC for enhancing the prophylactic or therapeutic respiratory effect of an
CC antiinflammatory steroid in a subject, for reducing or depleting levels
CC of, or reducing sensitivity to adenosine, reducing levels of adenosine
CC receptor, producing bronchodilation, increasing levels of ubiquinone or
CC lung surfactant in a subject's tissue, or treating bronchoconstriction,
CC lung inflammation, lung allergies, or a respiratory disease or condition.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 20 BP; 2 A; 13 C; 1 G; 4 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	10	Gaps:	0

US-10-031-158-14 (1-58) x ABZ85409 (1-20)

```
Qy          4 PheProProSerProLeu 9
              ||||||||||||||||||
Db          3 TTCCCTCCCAGCCCCCTC 20
```

RESULT 14
ABD21639
ID ABD21639 standard; DNA; 20 BP.
XX
AC ABD21639;
XX
DT 29-JUL-2004 (first entry)
XX
DE S100 calcium binding protein A2-derived oligo SEQ ID 651.
XX
KW Human; antisense; bronchoconstriction; allergy; hyposecretion; pain;
KW respiratory tract inflammation; adenosine sensitivity; lung; cancer;
KW surfactant depletion; antiallergic; antiinflammatory; antiasthmatic;
KW analgesic; hypotensive; immunosuppressive; cytostatic; cystic fibrosis;
KW beta-adrenergic agonist; respiratory disease; pulmonary vasoconstriction;
KW respiratory distress syndrome; allergic rhinitis; pulmonary hypertension;
KW emphysema; chronic obstructive pulmonary disease; cancer; bronchitis;
KW pulmonary transplantation rejection; ss; primer.
XX
OS Homo sapiens.
XX
PN WO200285309-A2.
XX
PD 31-OCT-2002.
XX
PF 23-APR-2002; 2002WO-US013143.
XX
PR 24-APR-2001; 2001US-0286036P.
XX
PA (EPIG-) EPIGENESIS PHARM INC.
XX
PI Nyce JW, Li Y, Sandrasagra A, Katz E, Pabalan J, Aguilar D;
PI Miller S, Tang L, Shahabuddin S;
XX
DR WPI; 2003-093058/08.
XX
PT Pharmaceutical composition for treating asthma, has antisense
PT oligonucleotide containing less percentage of adenosine, targeted to
PT nucleic acids associated with lung airway or lung dysfunction, and
PT bronchodilating agent.
XX
PS Claim 15; SEQ ID NO 651; 763pp; English.
XX
CC This invention describes a novel composition (a) a first active agent,
CC comprising oligonucleotides, effective for alleviating
CC bronchoconstriction, respiratory tract inflammation, allergies and
CC reducing adenosine sensitivity, levels of adenosine (A) or (A) receptors,
CC surfactant depletion or hyposecretion, when administered to a mammal. The
CC oligonucleotides are derived from a gene encoding or regulating
CC expression of a target polypeptide associated with lung airway or lung
CC dysfunction or cancer and can be anti-sense to the corresponding mRNA.
CC The invention also describes a kit, that comprises: (a) a delivery
CC device, in separate containers, (b) the oligonucleotides, (c)
CC instructions for adding a carrier and for use of the kit. The composition
CC of the invention has antiallergic, antiinflammatory, antiasthmatic,
CC analgesic, hypotensive, immunosuppressive and cytostatic activity, is a
CC beta-adrenergic agonist. The composition is useful for preventing or
CC treating a respiratory, lung or malignant disease. The administered
CC composition comprises oligo and is administered to reduce the production
CC or availability, or to increase the degradation of the target mRNA or to
CC reduce the amount of target polypeptide present in the lungs. The
CC pulmonary obstruction, and/or bronchoconstriction and/or lung
CC inflammation, allergies and/or surfactant hypoproduction are associated
CC with a disease or condition such as pulmonary vasoconstriction,
CC inflammation, allergies, asthma, impeded respiration, respiratory
CC distress syndrome, pain, cystic fibrosis, allergic rhinitis, pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease, pulmonary
CC transplantation rejection, pulmonary infections, bronchitis or cancer.
CC The reduced adenosine content of the anti-sense oligos corresponding to
CC thymidines present in the target RNA serves to prevent the breakdown of
CC the oligonucleotides into products that free adenosine into the system
CC e.g., lung, brain, heart, kidney, etc, tissue environment and thereby, to
CC prevent any unwanted effects due to it
XX
SQ Sequence 20 BP; 2 A; 13 C; 1 G; 4 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	11	Gaps:	0

US-10-031-158-14 (1-58) x ABD21639 (1-20)

```
Qy          4 PheProProSerProLeu 9
              ||||||||||||||||||
Db          3 TTCCCTCCCAGCCCCCTC 20
```

RESULT 15
ADS89854/c
ID ADS89854 standard; DNA; 20 BP.
XX
AC ADS89854;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human PCR primer SEQ ID NO:870.
XX
KW ss; PCR; primer; human; cell proliferative disorder; breast; methylation;
KW cytostatic; gene therapy; single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
XX
PN WO2004035803-A2.
XX
PD 29-APR-2004.
XX
PF 01-OCT-2003; 2003WO-EP010881.
XX
PR 01-OCT-2002; 2002DE-01045779.
PR 07-JAN-2003; 2003DE-01000096.
PR 17-APR-2003; 2003DE-01017955.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
PI Nimmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;
XX

DR WPI; 2004-348468/32.
XX
PT Predicting responsiveness of a subject with breast cell proliferative
PT disorder, useful for treating or differentiating breast cell
PT proliferative disorders comprises analyzing methylation pattern of a
PT genomic DNA from the subject.
XX
PS Example 2; SEQ ID NO 870; 104pp; English.
XX
CC The invention relates to a novel method for predicting the responsiveness
CC of a subject with a cell proliferative disorder of the breast tissues to
CC a therapy comprising analysing the methylation pattern of a target
CC nucleic acid by contacting at least one of the target nucleic acids in a
CC biological sample obtained from the subject prior to or during treatment.
CC The method of the invention has cytostatic activity, and may have a use
CC in gene therapy. The set of oligonucleotides comprising at least two of
CC the oligomers are useful for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
CC methods, nucleic acid, oligonucleotide, and kit are useful for the
CC treatment, characterisation, classification and/or differentiation, of
CC breast cell proliferative disorders. The method is also useful for
CC predicting the responsiveness of a subject with a cell proliferative
CC disorder of the breast tissues to a therapy. The present sequence is used
CC in the exemplification of the invention.
XX
SQ Sequence 20 BP; 8 A; 0 C; 11 G; 1 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	749	Length:	20
Score:	6.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.34%	Indels:	0
DB:	13	Gaps:	0

US-10-031-158-14 (1-58) x ADS89854 (1-20)

```
Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1
```

Search completed: October 16, 2005, 09:40:49
Job time : 423 secs

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 07:43:47 ; Search time 1775 Seconds
(without alignments)
1583.325 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1429811

Minimum DB seq length: 0
Maximum DB seq length: 30

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135842_19603/app_query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n_olig_sz30.rge -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=30
-USER=US10031158_@CGN_1_1_5600_@runat_14102005_135842_19603 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	8	13.8	27	6	AX074409	AX074409 Sequence
c	2	8	13.8	27	6	AX074410	AX074410 Sequence
c	3	7	12.1	23	6	BD263082	BD263082 Vector. 7
c	4	7	12.1	23	6	AX025028	AX025028 Sequence
	5	7	12.1	25	6	AX771994	AX771994 Sequence
c	6	6	10.3	20	6	CQ807420	CQ807420 Sequence
	7	6	10.3	20	6	AR299570	AR299570 Sequence
c	8	6	10.3	20	6	AX599128	AX599128 Sequence
	9	6	10.3	20	6	AX613466	AX613466 Sequence
c	10	6	10.3	20	6	AX796046	AX796046 Sequence
c	11	6	10.3	20	6	AX822579	AX822579 Sequence
c	12	6	10.3	20	6	AX826219	AX826219 Sequence
c	13	6	10.3	22	6	BD132408	BD132408 A basal c
	14	6	10.3	23	6	AX395174	AX395174 Sequence
	15	6	10.3	23	6	AX771993	AX771993 Sequence
	16	6	10.3	25	6	AX674984	AX674984 Sequence
	17	6	10.3	26	6	AX717020	AX717020 Sequence
	18	6	10.3	26	6	AX717041	AX717041 Sequence
c	19	6	10.3	30	6	AR000072	AR000072 Sequence
c	20	6	10.3	30	6	AR064895	AR064895 Sequence
c	21	6	10.3	30	6	AX793255	AX793255 Sequence
	22	6	10.3	30	9	HUMPLTP26	U37833 Human phosp
c	23	5	8.6	15	6	E49840	E49840 Novel phosp
c	24	5	8.6	15	6	BD083890	BD083890 Novel pho
c	25	5	8.6	15	6	BD096208	BD096208 Novel pho
c	26	5	8.6	16	6	AR002252	AR002252 Sequence
c	27	5	8.6	16	6	AR045202	AR045202 Sequence
c	28	5	8.6	16	6	AR051233	AR051233 Sequence
c	29	5	8.6	16	6	AR066236	AR066236 Sequence
c	30	5	8.6	16	6	AR101749	AR101749 Sequence
c	31	5	8.6	16	6	AR127780	AR127780 Sequence
	32	5	8.6	16	6	AR127795	AR127795 Sequence
c	33	5	8.6	16	6	AR166238	AR166238 Sequence
c	34	5	8.6	16	6	BD263078	BD263078 Vector. 7
c	35	5	8.6	16	6	BD273667	BD273667 Novel oli
c	36	5	8.6	16	6	I16027	I16027 Sequence 1
c	37	5	8.6	16	6	I28362	I28362 Sequence 1
c	38	5	8.6	16	6	I28568	I28568 Sequence 21
c	39	5	8.6	16	6	I28569	I28569 Sequence 22
c	40	5	8.6	16	6	I58730	I58730 Sequence 21
c	41	5	8.6	16	6	I58731	I58731 Sequence 22
c	42	5	8.6	16	6	AX025024	AX025024 Sequence
c	43	5	8.6	16	6	AX769906	AX769906 Sequence
c	44	5	8.6	17	6	A88667	A88667 Sequence 81
c	45	5	8.6	17	6	A90634	A90634 Sequence 81

ALIGNMENTS

RESULT 1
AX074409/c
LOCUS AX074409 27 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 7 from Patent WO0104309.
ACCESSION AX074409
VERSION AX074409.1 GI:12710547
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Pastan,I., Essand,M., Lee,B., Vasmatzis,G. and Wolfgang,C.
TITLE T-cell receptor _g(g) alternate reading frame protein, (tarp) and uses thereof
JOURNAL Patent: WO 0104309-A 7 18-JAN-2001;
UNITED STATES GOVERNMENT (US)
FEATURES Location/Qualifiers
source 1. .27
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="PCR primer"
ORIGIN

Alignment Scores:
Pred. No.: 3.79 Length: 27
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0

```
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            13.79%         Indels:         0
DB:                     6              Gaps:           0

US-10-031-158-14 (1-58) x AX074409 (1-27)

Qy         13 LeuGlnLeuLeuLysGlnSerSer 20
              |||||||||||||||||||||||
Db         26 CTTCAATTGCTGAAACAAAGCTCC 3

RESULT 2
AX074410/c
LOCUS       AX074410                  27 bp    DNA     linear   PAT 06-FEB-2001
DEFINITION  Sequence 8 from Patent WO0104309.
ACCESSION   AX074410
VERSION     AX074410.1  GI:12710548
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Pastan,I., Essand,M., Lee,B., Vasmatzis,G. and Wolfgang,C.
  TITLE     T-cell receptor _g(g) alternate reading frame protein, (tarp) and
            uses thereof
  JOURNAL   Patent: WO 0104309-A 8 18-JAN-2001;
            UNITED STATES GOVERNMENT (US)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PCR primer"
ORIGIN

Alignment Scores:
Pred. No.:              3.79           Length:         27
Score:                  8.00           Matches:        8
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            13.79%         Indels:         0
DB:                     6              Gaps:           0

US-10-031-158-14 (1-58) x AX074410 (1-27)

Qy         23 LeuGluHisThrPheValPheLeu 30
              |||||||||||||||||||||||
Db         26 CTGGAACATACCTTTGTCTTCTTG 3

RESULT 3
BD263082/c
LOCUS       BD263082                  23 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Vector.
ACCESSION   BD263082
VERSION     BD263082.1  GI:33072850
KEYWORDS    JP 2002530115-A/14.
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Mitrophanous,K., Uden,M., Rohll,J., Kingsman,S.M. and Kingsman,A.J.
  TITLE     Vector
  JOURNAL   Patent: JP 2002530115-A 14 17-SEP-2002;
            OXFORD BIOMEDICA LTD
COMMENT     OS   Human immunodeficiency virus type 1
            PN   JP 2002530115-A/14
            PD   17-SEP-2002
            PF   19-NOV-1999 JP 2000584089
            PR   20-NOV-1998 GB    9825524.3
            PI   KYRIACOS MITROPHANOUS,MARK UDEN,JONATHAN ROHLL,SUSAN MARY  PI
             KINGSMAN,
            PI   ALAN JOHN KINGSMAN
            PC   C12N15/09,A61K35/76,A61K48/00,A61P1/04,A61P9/00,A61P11/06, PC
             A61P17/00,
            PC   A61P25/00,A61P25/28,A61P27/02,A61P29/00,A61P31/12,A61P35/00,
            PC   A61P37/00,
            PC   C12N5/10,C12N7/00,C12N15/00,C12N5/00
            CC   Vector
            FH   Key             Location/Qualifiers
            FT   source          1. .23
            FT                   /organism='Human immunodeficiency virus type
            FT                   1'.
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:11676"
ORIGIN

Alignment Scores:
Pred. No.:              38             Length:         23
Score:                  7.00           Matches:        7
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            12.07%         Indels:         0
DB:                     6              Gaps:           0

US-10-031-158-14 (1-58) x BD263082 (1-23)

Qy          5 ProProSerProLeuPhePhe 11
              |||||||||||||||||||||
Db         22 CCCCCATCCCCCCTTTTCTTT 2

RESULT 4
AX025028/c
LOCUS       AX025028                  23 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 14 from Patent WO0031280.
ACCESSION   AX025028
VERSION     AX025028.1  GI:10184948
KEYWORDS    .
SOURCE      Human immunodeficiency virus 1 (HIV-1)
  ORGANISM  Human immunodeficiency virus 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1
  AUTHORS   Kingsman,S.M., Mitrophanous,K., Uden,M., Rohll,J. and Kingsman,A.J.
  TITLE     Vector
  JOURNAL   Patent: WO 0031280-A 14 02-JUN-2000;
            KINGSMAN SUSAN MARY (GB) ; MITROPHANOUS KYRIACOS (GB) ; UDEN MARK
            (GB) ; ROHLL JONATHAN (GB) ; KINGSMAN ALAN JOHN (GB) ; OXFORD
            BIOMEDICA LTD (GB)
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="Human immunodeficiency virus 1"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:11676"
ORIGIN

Alignment Scores:
Pred. No.:              38             Length:         23
Score:                  7.00           Matches:        7
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            12.07%         Indels:         0
DB:                     6              Gaps:           0

US-10-031-158-14 (1-58) x AX025028 (1-23)

Qy          5 ProProSerProLeuPhePhe 11
              |||||||||||||||||||||
Db         22 CCCCCATCCCCCCTTTTCTTT 2

RESULT 5
AX771994
LOCUS       AX771994                  25 bp    DNA     linear   PAT 02-JUL-2003
DEFINITION  Sequence 9 from Patent WO03006065.
```

```
ACCESSION   AX771994
VERSION     AX771994.1  GI:32438642
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Braun,K., Waldeck,W., Pipkorn,R., Braun,I. and Debus,J.
  TITLE     PNA conjugate for the treatment of diseases associated with HIV
  JOURNAL   Patent: WO 03006065-A 9 23-JAN-2003;
            Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
            (DE)
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="PNA Ia"
ORIGIN

Alignment Scores:
Pred. No.:                41.1          Length:          25
Score:                    7.00          Matches:         7
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              12.07%        Indels:          0
DB:                       6             Gaps:            0

US-10-031-158-14 (1-58) x AX771994 (1-25)

Qy          7 SerProLeuPhePhePheLeu 13
              |||||||||||||||||||||
Db          4 TCCCCCCTTTTCTTTTTTCTT 24

RESULT 6
CQ807420/c
LOCUS       CQ807420                  20 bp    DNA     linear   PAT 10-MAY-2004
DEFINITION  Sequence 870 from Patent WO2004035803.
ACCESSION   CQ807420
VERSION     CQ807420.1  GI:47112814
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Foekens,J., Harbeck,N., Koenig,T., Maier,S., Martens,J., Model,F.,
            Nimmrich,I., Rujan,T., Schmitt,A., Schmitt,M., Look,M.P. and
            Marx,A.
  TITLE     Method and nucleic acids for the improved treatment of breast cell
            proliferative disorders
  JOURNAL   Patent: WO 2004035803-A 870 29-APR-2004;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Detection primer for SFN"
ORIGIN

Alignment Scores:
Pred. No.:                389           Length:          20
Score:                    6.00          Matches:         6
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              10.34%        Indels:          0
DB:                       6             Gaps:            0

US-10-031-158-14 (1-58) x CQ807420 (1-20)

Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1

RESULT 7
AR299570
LOCUS       AR299570                  20 bp    DNA     linear   PAT 12-JUN-2003
DEFINITION  Sequence 11305 from patent US 6537751.
ACCESSION   AR299570
VERSION     AR299570.1  GI:31686854
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Cohen,D., Chumakov,I. and Blumenfeld,M.
  TITLE     Biallelic markers for use in constructing a high density
            disequilibrium map of the human genome
  JOURNAL   Patent: US 6537751-A 11305 25-MAR-2003;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Alignment Scores:
Pred. No.:                389           Length:          20
Score:                    6.00          Matches:         6
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              10.34%        Indels:          0
DB:                       6             Gaps:            0

US-10-031-158-14 (1-58) x AR299570 (1-20)

Qy          8 ProLeuPhePhePheLeu 13
              ||||||||||||||||||
Db          3 CCCCTATTTTTTTTCTG 20

RESULT 8
AX599128/c
LOCUS       AX599128                  20 bp    DNA     linear   PAT 14-FEB-2003
DEFINITION  Sequence 468 from Patent WO02077272.
ACCESSION   AX599128
VERSION     AX599128.1  GI:28399270
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Berlin,K., Braun,A., Distler,J., Guetig,D., Howe,A., Mueller,J.,
            Olek,A., Piepenbrock,C., Adorjan,P., Grabs,G., Lesche,R., Leu,E.,
            Lewin,A., Lipscher,E., Maier,S., Model,F., Mueller,V., Otto,T.,
            Pelet,C. and Ziebarth,H.
  TITLE     Methods and nucleic acids for the analysis of hematopoietic cell
            proliferative disorders
  JOURNAL   Patent: WO 02077272-A 468 03-OCT-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Detection primer for SFN"
ORIGIN

Alignment Scores:
Pred. No.:                389           Length:          20
Score:                    6.00          Matches:         6
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              10.34%        Indels:          0
DB:                       6             Gaps:            0

US-10-031-158-14 (1-58) x AX599128 (1-20)
```

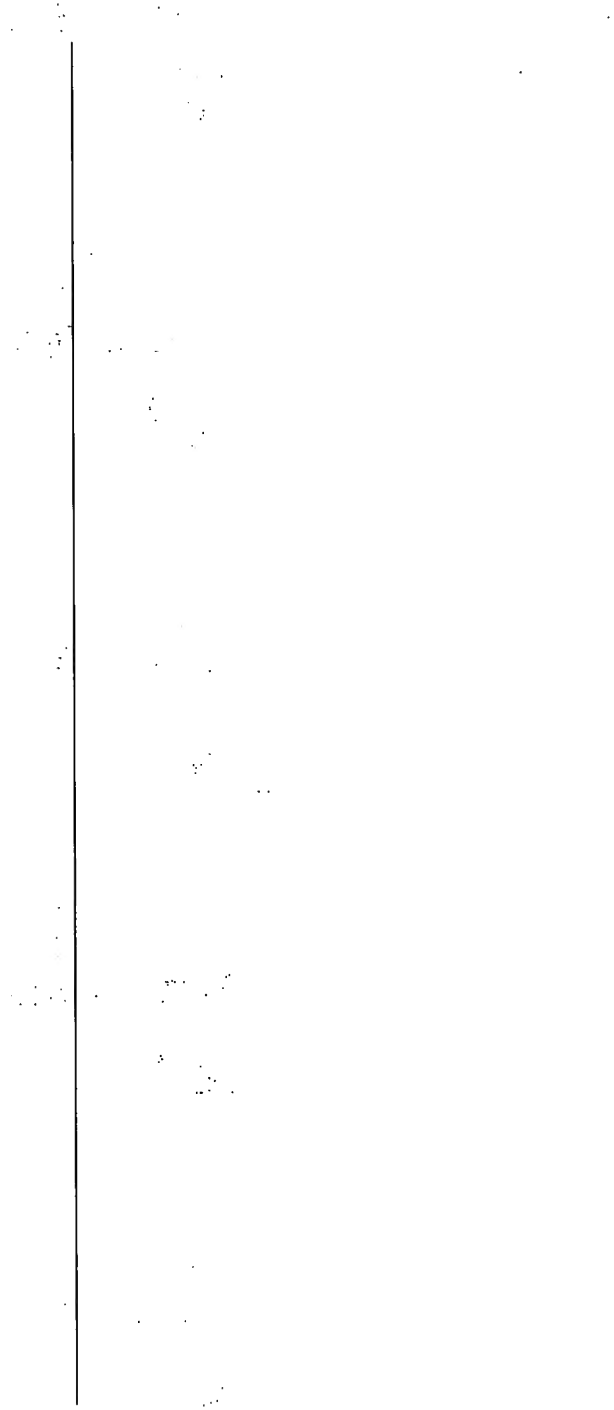
```
Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1

RESULT 9
AX613466
LOCUS       AX613466                  20 bp    DNA     linear   PAT 17-FEB-2003
DEFINITION  Sequence 4491 from Patent WO02072882.
ACCESSION   AX613466
VERSION     AX613466.1  GI:28408895
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Cullen,P. and Seedorf,U.
  TITLE     Coronary chip
  JOURNAL   Patent: WO 02072882-A 4491 19-SEP-2002;
            OGHAM GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              389           Length:         20
Score:                  6.00          Matches:        6
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            10.34%        Indels:         0
DB:                     6             Gaps:           0

US-10-031-158-14 (1-58) x AX613466 (1-20)

Qy         42 GlyLysLysArgArgAla 47
              ||||||||||||||||||
Db          1 GGAAAGAAACGCAGGGCT 18

RESULT 10
AX796046/c
LOCUS       AX796046                  20 bp    DNA     linear   PAT 04-OCT-2003
DEFINITION  Sequence 389 from Patent WO03052135.
ACCESSION   AX796046
VERSION     AX796046.1  GI:37516712
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Burger,M., Field,J.K., Genc,B., Liloglou,T., Lipscher,E., Maier,S.
            and Nimmrich,I.
  TITLE     Method and nucleic acids for the analysis of a lung cell
            proliferative disorder
  JOURNAL   Patent: WO 03052135-A 389 26-JUN-2003;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Detection primer for SFN"
ORIGIN

Alignment Scores:
Pred. No.:              389           Length:         20
Score:                  6.00          Matches:        6
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            10.34%        Indels:         0
DB:                     6             Gaps:           0

US-10-031-158-14 (1-58) x AX796046 (1-20)

Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1

RESULT 11
AX822579/c
LOCUS       AX822579                  20 bp    DNA     linear   PAT 11-DEC-2003
DEFINITION  Sequence 471 from Patent EP1340818.
ACCESSION   AX822579
VERSION     AX822579.1  GI:39749215
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Adorjan,P., Burger,M., Maier,S., Nimmrich,I., Becker,E., Lesche,R.,
            Rujan,T. and Schmitt,A.
  TITLE     Method and nucleic acids for the analysis of a colon cell
            proliferative disorder
  JOURNAL   Patent: EP 1340818-A 471 03-SEP-2003;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Detection primer for SFN"
ORIGIN

Alignment Scores:
Pred. No.:              389           Length:         20
Score:                  6.00          Matches:        6
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            10.34%        Indels:         0
DB:                     6             Gaps:           0

US-10-031-158-14 (1-58) x AX822579 (1-20)

Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1

RESULT 12
AX826219/c
LOCUS       AX826219                  20 bp    DNA     linear   PAT 11-DEC-2003
DEFINITION  Sequence 471 from Patent WO03072821.
ACCESSION   AX826219
VERSION     AX826219.1  GI:39751733
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Adorjan,P., Burger,M., Maier,S., Nimmrich,I., Becker,E., Lesche,R.,
            Rujan,T. and Schmitt,A.
  TITLE     Method and nucleic acids for the analysis of a colon cell
            proliferactive disorder
  JOURNAL   Patent: WO 03072821-A 471 04-SEP-2003;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Detection primer for SFN"
ORIGIN

Alignment Scores:
Pred. No.:              389           Length:         20
```

```
Score:                  6.00          Matches:       6
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            10.34%        Indels:        0
DB:                     6             Gaps:          0

US-10-031-158-14 (1-58) x AX826219 (1-20)

Qy          5 ProProSerProLeuPhe 10
              ||||||||||||||||||
Db         18 CCTCCCTCTCCTCTCTTC 1

RESULT 13
BD132408/c
LOCUS       BD132408                22 bp    DNA     linear   PAT 18-SEP-2002
DEFINITION  A basal cell carcinoma tumor suppressor gene.
ACCESSION   BD132408
VERSION     BD132408.1  GI:23227353
KEYWORDS    JP 2002504805-A/20.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Dean,M.F., Hahn,H., Wicking,C., Christiansen,J.,
            Zaphiropoulos,P.G., Gailani,M.R., Shanley,S., Chidambaram,A.,
            Vorechovsky,I., Holmberg,E., Unden,A.B., Gillies,S., Negus,K.,
            Smyth,I., Pressman,C., Leffell,D.J., Gerrard,B., Goldstein,A.,
            Wainwright,B., Toftgard,R., Trench,G.C. and Bale,A.E.
  TITLE     A basal cell carcinoma tumor suppressor gene
  JOURNAL   Patent: JP 2002504805-A 20 12-FEB-2002;
            THE GOVERNMENT OF THE UNITED STATES OF AMERICA REPRESENTED BY THE
            SECRETARY DEPARTMENT OF HEALTH AND HUMAN SERVICES
COMMENT     PN   JP 2002504805-A/20
            PD   12-FEB-2002
            PF   16-MAY-1997 JP 1997541164
            PR   17-MAY-1996 US    60/017906,21-MAY-1996 AU      PO 0011 PR
            07-JUN-1996 AU      PO 0363,14-JUN-1996 US    60/019765 PI
            MICHAEL FREDERICK DEAN,HEIDI HAHN,CAROL WICKING,JEFFREY  PI
            CHRISTIANSEN,
            PI   PETER G ZAPHIROPOULOS,MAE R GAILANI,SUSAN SHANLEY,ABIRAMI  PI
             CHIDAMBARAM,
            PI   IGOR VORECHOVSKY,ERIKA HOLMBERG,ANNE BIRGITTE UNDEN,SUSAN  PI
             GILLIES,
            PI   KYLIE NEGUS,IAN SMYTH,CAROL PRESSMAN,DAVID J LEFFELL,BERNARD
            PI   GERRARD,
            PI   ALISA GOLDSTEIN,BRANDON WAINWRIGHT,RUNE TOFTGARD,GEORGIA  PI
            CHENEVIX TRENCH,
            PI   ALLEN E BALE
            PC   C12N15/12,C07K14/47,C12N5/10,C12Q1/68,G01N33/50,A61K48/00, PC
             A61K39/395,
            PC   A61K38/17
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   /note= 'PTCF21 primer'
            FH   Key             Location/Qualifiers.
FEATURES            Location/Qualifiers
     source         1. .22
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
ORIGIN

Alignment Scores:
Pred. No.:              425           Length:        22
Score:                  6.00          Matches:       6
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            10.34%        Indels:        0
DB:                     6             Gaps:          0

US-10-031-158-14 (1-58) x BD132408 (1-22)

Qy          28 ValPheLeuArgAsnPhe 33
                 ||||||||||||||||||
Db          21 GTGTTCCTGAGAAATTTT 4

RESULT 14
AX395174
LOCUS       AX395174                23 bp    DNA     linear   PAT 18-MAY-2002
DEFINITION  Sequence 6 from Patent WO0218579.
ACCESSION   AX395174
VERSION     AX395174.1  GI:21066224
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Glucksmann,M.A.
  TITLE     57242, a novel human g protein-coupled receptor family member and
            uses therefor
  JOURNAL   Patent: WO 0218579-A 6 07-MAR-2002;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES            Location/Qualifiers
     source         1. .23
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
                    /note="murine 57242 probe sequence"
ORIGIN

Alignment Scores:
Pred. No.:              443           Length:        23
Score:                  6.00          Matches:       6
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            10.34%        Indels:        0
DB:                     6             Gaps:          0

US-10-031-158-14 (1-58) x AX395174 (1-23)

Qy          45 ArgArgAlaThrArgPhe 50
                 ||||||||||||||||||
Db           4 AGGAGGGCCACCCGGTTC 21

RESULT 15
AX771993
LOCUS       AX771993                23 bp    DNA     linear   PAT 02-JUL-2003
DEFINITION  Sequence 8 from Patent WO03006065.
ACCESSION   AX771993
VERSION     AX771993.1  GI:32438641
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Braun,K., Waldeck,W., Pipkorn,R., Braun,I. and Debus,J.
  TITLE     PNA conjugate for the treatment of diseases associated with HIV
  JOURNAL   Patent: WO 03006065-A 8 23-JAN-2003;
            Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts
            (DE)
FEATURES            Location/Qualifiers
     source         1. .23
                    /organism="synthetic construct"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:32630"
                    /note="PNA I"
ORIGIN

Alignment Scores:
Pred. No.:              443           Length:        23
Score:                  6.00          Matches:       6
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            10.34%        Indels:        0
DB:                     6             Gaps:          0
```

US-10-031-158-14 (1-58) x AX771993 (1-23)

```
Qy          7 SerProLeuPhePhePhe 12
              ||||||||||||||||||
Db          1 TCCCCCCTTTTCTTTTTT 18
```

Search completed: October 16, 2005, 10:10:35
Job time : 1777 secs



GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:59:35 ; Search time 113 Seconds
(without alignments)
213.891 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1859788 seqs, 416717961 residues

Word size : 0

Total number of hits satisfying chosen parameters: 230707

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	8.6	5	9	US-09-895-443-30	Sequence 30, Appl
2	5	8.6	7	9	US-09-833-067-26	Sequence 26, Appl
3	5	8.6	8	15	US-10-137-867-292	Sequence 292, App
4	5	8.6	9	9	US-09-765-527-221	Sequence 221, App
5	5	8.6	9	9	US-09-792-480-18	Sequence 18, Appl
6	5	8.6	9	11	US-09-870-216C-3	Sequence 3, Appli
7	5	8.6	9	13	US-10-017-327-3	Sequence 3, Appli
8	5	8.6	9	14	US-10-144-188-2	Sequence 2, Appli
9	5	8.6	10	9	US-09-839-329-4	Sequence 4, Appli
10	5	8.6	10	10	US-09-572-404B-238	Sequence 238, App
11	5	8.6	10	10	US-09-572-404B-1444	Sequence 1444, Ap
12	5	8.6	10	10	US-09-572-404B-3212	Sequence 3212, Ap
13	5	8.6	10	13	US-10-007-761-9	Sequence 9, Appli
14	5	8.6	10	14	US-10-083-960-18	Sequence 18, Appl
15	5	8.6	10	14	US-10-209-421-18	Sequence 18, Appl
16	5	8.6	10	14	US-10-168-789A-26	Sequence 26, Appl
17	5	8.6	10	14	US-10-168-789A-27	Sequence 27, Appl
18	5	8.6	10	15	US-10-421-548-7	Sequence 7, Appli
19	5	8.6	10	15	US-10-421-503-65	Sequence 65, Appl
20	5	8.6	10	16	US-10-843-731-9	Sequence 9, Appli
21	5	8.6	10	17	US-10-936-237-47	Sequence 47, Appl
22	5	8.6	10	18	US-10-513-003-65	Sequence 65, Appl
23	4	6.9	4	10	US-09-793-451-717	Sequence 717, App
24	4	6.9	4	14	US-10-062-109A-701	Sequence 701, App
25	4	6.9	4	14	US-10-005-480A-701	Sequence 701, App
26	4	6.9	4	14	US-10-283-722-717	Sequence 717, App
27	4	6.9	4	15	US-10-291-241-41	Sequence 41, Appl
28	4	6.9	4	15	US-10-283-903-717	Sequence 717, App
29	4	6.9	4	15	US-10-133-234A-9	Sequence 9, Appli
30	4	6.9	4	17	US-10-418-032-253	Sequence 253, App
31	4	6.9	4	18	US-10-892-402-36	Sequence 36, Appl
32	4	6.9	5	9	US-09-895-443-24	Sequence 24, Appl
33	4	6.9	5	9	US-09-895-443-25	Sequence 25, Appl
34	4	6.9	5	9	US-09-895-443-26	Sequence 26, Appl
35	4	6.9	5	9	US-09-781-133-20	Sequence 20, Appl
36	4	6.9	5	9	US-09-781-133-21	Sequence 21, Appl
37	4	6.9	5	9	US-09-781-133-22	Sequence 22, Appl
38	4	6.9	5	14	US-10-235-483-29	Sequence 29, Appl
39	4	6.9	5	14	US-10-299-540-2	Sequence 2, Appli
40	4	6.9	5	14	US-10-299-184-2	Sequence 2, Appli
41	4	6.9	5	16	US-10-703-206-3	Sequence 3, Appli
42	4	6.9	5	17	US-10-808-187-1075	Sequence 1075, Ap
43	4	6.9	5	17	US-10-497-628-5	Sequence 5, Appli
44	4	6.9	5	17	US-10-776-224-204	Sequence 204, App
45	4	6.9	5	18	US-10-807-807-1075	Sequence 1075, Ap

ALIGNMENTS

RESULT 1
US-09-895-443-30
; Sequence 30, Application US/09895443
; Patent No. US20020103134A1
; GENERAL INFORMATION:
; -APPLICANT: Findeis, M. et al.
; TITLE OF INVENTION: Modulators of -Amyloid Peptide
; Aggregation Comprising D-Amino Acids
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1784
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/895,443
; FILING DATE: 29-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/356,931
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/897,342
; FILING DATE: 21-JUL-1997
; APPLICATION NUMBER: USSN 08/703,675
; FILING DATE: 27-AUG-1996
; APPLICATION NUMBER: USSN 08/616,081
; FILING DATE: 14-MAR-1996

```
;             APPLICATION NUMBER: USSN 08/548,998
;             FILING DATE: 27-OCT-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: DeConti, Jr., Giulio A.
;             REGISTRATION NUMBER: 31,503
;             REFERENCE/DOCKET NUMBER: PPI-016CP4CN
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)742-4214
;   INFORMATION FOR SEQ ID NO: 30:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 5 amino acids
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  Modified site
;             LOCATION:  1-5
;             OTHER INFORMATION:  /note= D amino acid
;        SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-09-895-443-30

  Query Match             8.6%;  Score 5;  DB 9;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 2
US-09-833-067-26
; Sequence 26, Application US/09833067
; Patent No. US20020054888A1
; GENERAL INFORMATION:
;  APPLICANT: O'HANLEY, PETER
;  APPLICANT:  DENICH, KENNETH
;  TITLE OF INVENTION: DISSOCIATED PILI, THEIR PRODUCTION AND USE
;  FILE REFERENCE: 050939/0102
;  CURRENT APPLICATION NUMBER: US/09/833,067
;  CURRENT FILING DATE:  2001-04-12
;  PRIOR APPLICATION NUMBER: 60/196,493
;  PRIOR FILING DATE: 2000-04-12
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Unknown Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Sequence: highly hydrophobic region
;   OTHER INFORMATION: of amino acids
US-09-833-067-26

  Query Match             8.6%;  Score 5;  DB 9;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 PLFFF 12
              |||||
Db          3 PLFFF 7


RESULT 3
US-10-137-867-292
; Sequence 292, Application US/10137867
; Publication No. US20030207349A1
; GENERAL INFORMATION:
;  APPLICANT: Baker,Kevin P.
;  APPLICANT:  Beresini,Maureen
;  APPLICANT:  DeForge,Laura
;  APPLICANT:  Desnoyers,Luc
;  APPLICANT:  Filvaroff,Ellen
;  APPLICANT:  Gao,Wei-Qiang
;  APPLICANT:  Gerritsen,Mary E.
;  APPLICANT:  Goddard,Audrey
;  APPLICANT:  Godowski,Paul J.
;  APPLICANT:  Gurney,Austin L.
;  APPLICANT:  Sherwood,Steven
;  APPLICANT:  Smith,Victoria
;  APPLICANT:  Stewart,Timothy A.
;  APPLICANT:  Tumas,Daniel
;  APPLICANT:  Watanabe,Colin K
;  APPLICANT:  Wood,William
;  APPLICANT:  Zhang, Zemin
;  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;  TITLE OF INVENTION:  ACIDS ENCODING THE SAME
;  FILE REFERENCE: P3330R1C146
;  CURRENT APPLICATION NUMBER: US/10/137,867
;  CURRENT FILING DATE:  2002-05-03
;  Prior Application removed - See Palm or File Wrapper
;  NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 292
;   LENGTH: 428
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-137-867-292

  Query Match             8.6%;  Score 5;  DB 15;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 4
US-09-765-527-221
; Sequence 221, Application US/09765527
; Patent No. US20020006638A1
;   GENERAL INFORMATION:
;        APPLICANT: Better, Marc D.
;        TITLE OF INVENTION: Methods for Recombinant Microbial Production of
;                            Fusion Proteins and BPI-Derived Peptides
;        NUMBER OF SEQUENCES: 265
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;             STREET: 6300 Sears Tower, 233 South Wacker Drive
;             CITY: Chicago
;             STATE: Illinois
;             COUNTRY: United States of America
;             ZIP: 60606-6402
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,527
;             FILING DATE: 18-Jan-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/621,803
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Borun, Michael F.
;             REGISTRATION NUMBER: 25,447
;             REFERENCE/DOCKET NUMBER: 27129/33199
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 312/474-6300
;             TELEFAX: 312/474-0448
;             TELEX: 25-3856
;   INFORMATION FOR SEQ ID NO: 221:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 9 amino acids
```

```
;             TYPE: amino acid
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        FEATURE:
;             NAME/KEY:  misc_feature
;             OTHER INFORMATION:  "XMP.382:
;        FEATURE:
;             NAME/KEY:  Modified-site
;             LOCATION:  C-Terminus
;             OTHER INFORMATION:  /label= Amidation
;             /note= "The C-Terminus is Amidated."
;        SEQUENCE DESCRIPTION: SEQ ID NO: 221:
US-09-765-527-221

  Query Match             8.6%;  Score 5;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 LQLLK 17
              |||||
Db          4 LQLLK 8


RESULT 5
US-09-792-480-18
; Sequence 18, Application US/09792480
; Patent No. US20020127198A1
; GENERAL INFORMATION:
;  APPLICANT: Rothbard, Jonathan B.
;  APPLICANT:  Wender, Paul A.
;  APPLICANT:  McGrane, P. Leo
;  APPLICANT:  Sista, Lalitha V.S.
;  APPLICANT:  Kirschberg, Thorsten A.
;  APPLICANT:  CellGate, Inc.
;  TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
;  TITLE OF INVENTION:  Across and Into Epithelial Tissues
;  FILE REFERENCE: 019801-000230US
;  CURRENT APPLICATION NUMBER: US/09/792,480
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: US 09/648,400
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/150,510
;  PRIOR FILING DATE: 1999-08-24
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:A-54
;   OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
;   OTHER INFORMATION: region Tat-49-57
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: Xaa = fluorescein linked to amino group of
;   OTHER INFORMATION: aminohexanoic acid (Fl-ahx) attached to the
;   OTHER INFORMATION: N-terminal amino group of Arg
US-09-792-480-18

  Query Match             8.6%;  Score 5;  DB 9;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRRA 47
              |||||
Db          2 KKRRA 6


RESULT 6
US-09-870-216C-3
; Sequence 3, Application US/09870216C
; Publication No. US20040138135A1
; GENERAL INFORMATION:
;  APPLICANT: Charles A. Nicolette
;  TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER
;  FILE REFERENCE: 68126881210100
;  CURRENT APPLICATION NUMBER: US/09/870,216C
;  CURRENT FILING DATE:  2001-05-30
;  PRIOR APPLICATION NUMBER: 60/209,391
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: 60/226,256
;  PRIOR FILING DATE: 2000-08-17
;  PRIOR APPLICATION NUMBER: 60/257,008
;  PRIOR FILING DATE: 2000-12-20
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-870-216C-3

  Query Match             8.6%;  Score 5;  DB 11;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 FLQLL 16
              |||||
Db          1 FLQLL 5


RESULT 7
US-10-017-327-3
; Sequence 3, Application US/10017327
; Publication No. US20020155471A1
; GENERAL INFORMATION:
;  APPLICANT: Charles A. Nicolette
;  TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND
;  TITLE OF INVENTION:  METHODS FOR USING SAME
;  FILE REFERENCE: GZ 2101.20
;  CURRENT APPLICATION NUMBER: US/10/017,327
;  CURRENT FILING DATE:  2001-12-06
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-017-327-3

  Query Match             8.6%;  Score 5;  DB 13;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 FLQLL 16
              |||||
Db          1 FLQLL 5


RESULT 8
US-10-144-188-2
; Sequence 2, Application US/10144188
; Publication No. US20030170212A1
; GENERAL INFORMATION:
;  APPLICANT: Cai, Zeling
;  APPLICANT:  Jackson, Michael R.
;  APPLICANT:  Peterson, Per A.
;  APPLICANT:  Shi, Weixing
;  APPLICANT:  Kong, Yan
;  APPLICANT:  Degraw, Juli
;  TITLE OF INVENTION: Ex-Vivo Priming For Generating Cytotoxic T Lymphocytes Specific
;  TITLE OF INVENTION:  For No. US20030170212A1-Tumor Antigens To Treat Autoimmune And
;  FILE REFERENCE: PRI0010  ORT-1627
;  CURRENT APPLICATION NUMBER: US/10/144,188
;  CURRENT FILING DATE:  2002-05-13
```

```
;  PRIOR APPLICATION NUMBER: 60/291,300
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 60
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Peptide antigen
US-10-144-188-2

  Query Match             8.6%;  Score 5;  DB 14;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 PPSPL 9
              |||||
Db          2 PPSPL 6


RESULT 9
US-09-839-329-4
; Sequence 4, Application US/09839329
; Publication No. US20020022027A1
; GENERAL INFORMATION:
;  APPLICANT: Michael P. Sherman
;  APPLICANT:  Warner C. Greene
;  APPLICANT:  Carlos M.C. de No. 6664040ohna
;  APPLICANT:  Ulrich Schubert
;  APPLICANT:  Peter Henklein
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF
;  TITLE OF INVENTION:  A MOLECULE INTO A CELL
;  FILE REFERENCE: G&C 30448.91-US-U2
;  CURRENT APPLICATION NUMBER: US/09/839,329
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: 60/206,610
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/267,827
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Human immunodeficiency virus
US-09-839-329-4

  Query Match             8.6%;  Score 5;  DB 9;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         42 GKKRR 46
              |||||
Db          2 GKKRR 6


RESULT 10
US-09-572-404B-238
; Sequence 238, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 238
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in MTNR1B at 351-360 and may interact with Seque
;   OTHER INFORMATION: in this patent.
US-09-572-404B-238

  Query Match             8.6%;  Score 5;  DB 10;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         20 SRRLE 24
              |||||
Db          1 SRRLE 5


RESULT 11
US-09-572-404B-1444
; Sequence 1444, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 1444
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in NTAK at 681-690 and may interact with Sequenc
;   OTHER INFORMATION: in this patent.
US-09-572-404B-1444

  Query Match             8.6%;  Score 5;  DB 10;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         53 PRRGT 57
              |||||
Db          6 PRRGT 10


RESULT 12
US-09-572-404B-3212
; Sequence 3212, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
;  APPLICANT: Proteom Ltd
;  TITLE OF INVENTION: Complementary peptide ligands from the human genome
;  FILE REFERENCE: Human patent
;  CURRENT APPLICATION NUMBER: US/09/572,404B
;  CURRENT FILING DATE:  2000-05-17
;  NUMBER OF SEQ ID NOS: 4203
;  SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 3212
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   OTHER INFORMATION: sequence located in HOXA9 OR HOX1G at 369-378 and may interact wi
;   OTHER INFORMATION: Sequence 3211 in this patent.
US-09-572-404B-3212

  Query Match             8.6%;  Score 5;  DB 10;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 RRATR 49
              |||||
```

```
Db          4 RRATR 8


RESULT 13
US-10-007-761-9
; Sequence 9, Application US/10007761
; Publication No. US20020150984A1
; GENERAL INFORMATION:
;  APPLICANT: Mochly-Rosen, Daria
;  TITLE OF INVENTION: Peptides for Activation and Inhibition
;  TITLE OF INVENTION:  of delta-PKC
;  FILE REFERENCE: 58600-8208.US00
;  CURRENT APPLICATION NUMBER: US/10/007,761
;  CURRENT FILING DATE:  2001-11-09
;  PRIOR APPLICATION NUMBER: US 60/262,060
;  PRIOR FILING DATE: 2001-01-18
;  NUMBER OF SEQ ID NOS: 72
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Tat-derived carrier peptide
US-10-007-761-9

  Query Match             8.6%;  Score 5;  DB 13;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         42 GKKRR 46
              |||||
Db          2 GKKRR 6


RESULT 14
US-10-083-960-18
; Sequence 18, Application US/10083960
; Publication No. US20030022831A1
; GENERAL INFORMATION:
;  APPLICANT: Rothbard, Jonathan B.
;  APPLICANT:  Wender, Paul A.
;  APPLICANT:  McGrane, P. Leo
;  APPLICANT:  Sista, Lalitha V.S.
;  APPLICANT:  Kirschberg, Thorsten A.
;  APPLICANT:  CellGate, Inc.
;  TITLE OF INVENTION: Compositions and Methods for Enhancing
;  TITLE OF INVENTION:  Drug Delivery Across and Into Ocular Tissues
;  FILE REFERENCE: 019801-000240US
;  CURRENT APPLICATION NUMBER: US/10/083,960
;  CURRENT FILING DATE:  2003-07-14
;  PRIOR APPLICATION NUMBER: US 60/150,510
;  PRIOR FILING DATE: 1999-08-24
;  PRIOR APPLICATION NUMBER: US 09/648,400
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 09/792,480
;  PRIOR FILING DATE: 2001-02-23
;  NUMBER OF SEQ ID NOS: 86
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: A-54 alanine-substituted analog of Tat-49-57
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (1)...(1)
;   OTHER INFORMATION: Xaa = fluorescein conjugated aminohexanoic acid
;   OTHER INFORMATION: (Fl-ahx)
US-10-083-960-18

  Query Match             8.6%;  Score 5;  DB 14;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRRA 47
              |||||
Db          3 KKRRA 7


RESULT 15
US-10-209-421-18
; Sequence 18, Application US/10209421
; Publication No. US20030083256A1
; GENERAL INFORMATION:
;  APPLICANT: Rothbard, Jonathan B.
;  APPLICANT:  Wender, Paul A.
;  APPLICANT:  McGrane, P. Leo
;  APPLICANT:  Sista, Lalitha V.S.
;  APPLICANT:  Kirschberg, Thorsten A.
;  APPLICANT:  CellGate, Inc.
;  TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
;  TITLE OF INVENTION:  Across and Into Epithelial Tissues
;  FILE REFERENCE: 019801-000211US
;  CURRENT APPLICATION NUMBER: US/10/209,421
;  CURRENT FILING DATE:  2002-07-30
;  PRIOR APPLICATION NUMBER: US 60/150,510
;  PRIOR FILING DATE: 1999-08-24
;  PRIOR APPLICATION NUMBER: US 09/648,400
;  PRIOR FILING DATE: 2000-08-24
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:A-54
;   OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
;   OTHER INFORMATION: region Tat-49-57
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: Xaa = fluorescein linked to amino group of
;   OTHER INFORMATION: aminohexanoic acid (Fl-ahx)
US-10-209-421-18

  Query Match             8.6%;  Score 5;  DB 14;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRRA 47
              |||||
Db          3 KKRRA 7


Search completed: October 16, 2005, 06:11:00
Job time : 114 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:53:10 ; Search time 55 Seconds
(without alignments)
78.721 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 115750

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	8.6	5	2	US-08-920-162A-30	Sequence 30, Appl
2	5	8.6	5	3	US-09-356-931-30	Sequence 30, Appl
3	5	8.6	5	4	US-09-895-443A-30	Sequence 30, Appl
4	5	8.6	7	4	US-09-627-851B-47	Sequence 47, Appl
5	5	8.6	9	2	US-08-621-803-221	Sequence 221, App
6	5	8.6	9	2	US-08-621-259A-213	Sequence 213, App
7	5	8.6	9	3	US-08-159-339A-343	Sequence 343, App
8	5	8.6	9	3	US-08-159-339A-405	Sequence 405, App
9	5	8.6	9	3	US-09-217-352-221	Sequence 221, App
10	5	8.6	9	4	US-09-677-664B-213	Sequence 213, App
11	5	8.6	9	4	US-09-792-480-18	Sequence 18, Appl
12	5	8.6	10	2	US-08-248-839C-135	Sequence 135, App
13	5	8.6	10	4	US-09-648-400A-18	Sequence 18, Appl
14	5	8.6	10	4	US-09-839-329-4	Sequence 4, Appli
15	5	8.6	10	4	US-10-209-421-18	Sequence 18, Appl
16	5	8.6	10	5	PCT-US96-01735-4	Sequence 4, Appli
17	4	6.9	4	1	US-07-917-034A-3	Sequence 3, Appli
18	4	6.9	4	3	US-09-085-305-29	Sequence 29, Appl
19	4	6.9	4	4	US-09-627-851B-21	Sequence 21, Appl
20	4	6.9	4	4	US-09-627-851B-29	Sequence 29, Appl
21	4	6.9	4	5	PCT-US93-01669-36	Sequence 36, Appl
22	4	6.9	5	1	US-07-690-284A-2	Sequence 2, Appli
23	4	6.9	5	1	US-07-690-284A-5	Sequence 5, Appli
24	4	6.9	5	1	US-08-704-170-91	Sequence 91, Appl
25	4	6.9	5	1	US-08-618-464-2	Sequence 2, Appli
26	4	6.9	5	2	US-08-656-906-28	Sequence 28, Appl
27	4	6.9	5	2	US-08-920-162A-24	Sequence 24, Appl
28	4	6.9	5	2	US-08-920-162A-25	Sequence 25, Appl
29	4	6.9	5	2	US-08-920-162A-26	Sequence 26, Appl
30	4	6.9	5	3	US-08-752-526-13	Sequence 13, Appl
31	4	6.9	5	3	US-09-107-615-2	Sequence 2, Appli
32	4	6.9	5	3	US-09-217-847-28	Sequence 28, Appl
33	4	6.9	5	3	US-09-356-931-24	Sequence 24, Appl
34	4	6.9	5	3	US-09-356-931-25	Sequence 25, Appl
35	4	6.9	5	3	US-09-356-931-26	Sequence 26, Appl
36	4	6.9	5	3	US-08-703-675C-24	Sequence 24, Appl
37	4	6.9	5	3	US-08-703-675C-25	Sequence 25, Appl
38	4	6.9	5	3	US-08-703-675C-26	Sequence 26, Appl
39	4	6.9	5	4	US-08-766-596A-29	Sequence 29, Appl
40	4	6.9	5	4	US-09-895-443A-24	Sequence 24, Appl
41	4	6.9	5	4	US-09-895-443A-25	Sequence 25, Appl
42	4	6.9	5	4	US-09-895-443A-26	Sequence 26, Appl
43	4	6.9	5	4	US-09-674-973A-204	Sequence 204, App
44	4	6.9	5	5	PCT-US94-02631-91	Sequence 91, Appl
45	4	6.9	5	5	PCT-US95-04896-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-920-162A-30
; Sequence 30, Application US/08920162A
; Patent No. 5985242
;  GENERAL INFORMATION:
;    APPLICANT:  Findeis, M. et al.
;    TITLE OF INVENTION:  Modulators of {SYMBOL 98 \f "Symbol"}-Amyloid Peptide
;    TITLE OF INVENTION:  Aggregation Comprising D-Amino Acids
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/920,162A
;      FILING DATE:  Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/548,998
;      FILING DATE:  27-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/616,081
;      FILING DATE:  14-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/703,675
;      FILING DATE:  27-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:  21-JUL-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KARA, Catherine J.
;      REGISTRATION NUMBER:  41,106
;      REFERENCE/DOCKET NUMBER:  PPI-016CP4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
```

```
;    FEATURE:
;      NAME/KEY:  Modified site
;      LOCATION:  1-5
;      OTHER INFORMATION:  /note= D amino acid
US-08-920-162A-30

  Query Match             8.6%;  Score 5;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 2
US-09-356-931-30
; Sequence 30, Application US/09356931
; Patent No. 6277826
;  GENERAL INFORMATION:
;    APPLICANT:  Findeis, M. et al.
;    TITLE OF INVENTION:  Modulators of -Amyloid Peptide
;    TITLE OF INVENTION:  Aggregation Comprising D-Amino Acids
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1784
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/356,931
;      FILING DATE:  19-JUL-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/920,162
;      FILING DATE:  27-AUG-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/897,342
;      FILING DATE:  21-JUL-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/703,675
;      FILING DATE:  27-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/616,081
;      FILING DATE:  14-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/548,998
;      FILING DATE:  27-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeConti, Jr., Giulio A.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  PPI-016CP4CN
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)742-4214
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified site
;      LOCATION:  1-5
;      OTHER INFORMATION:  /note= D amino acid
US-09-356-931-30

  Query Match             8.6%;  Score 5;  DB 3;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 3
US-09-895-443A-30
; Sequence 30, Application US/09895443A
; Patent No. 6689752
; GENERAL INFORMATION:
;  APPLICANT: FINDEIS, MARK A.
;  APPLICANT:  GEFTER, MALCOLM L.
;  APPLICANT:  MUSSO, GARY
;  APPLICANT:  SIGNER, ETHAN R.
;  APPLICANT:  WAKEFIELD, JAMES
;  APPLICANT:  MOLINAUX, SUSAN
;  APPLICANT:  CHIN, JOSEPH
;  APPLICANT:  LEE, JUNG-JA
;  APPLICANT:  KELLEY, MICHAEL
;  APPLICANT:  KOMAR-PANICUCCI, SONJA
;  APPLICANT:  ARICO-MUENDEL, CHRISTOPHER C.
;  APPLICANT:  PHILLIPS, KATHRYN
;  APPLICANT:  HAYWARD, NEIL J.
;  TITLE OF INVENTION: MODULATORS OF BETA-AMYLOID PEPTIDE
;  TITLE OF INVENTION:  AGGREGATION COMPRISING D-AMINO ACIDS
;  FILE REFERENCE: PPI-0016CP4CN2
;  CURRENT APPLICATION NUMBER: US/09/895,443A
;  CURRENT FILING DATE:  2001-06-29
;  PRIOR APPLICATION NUMBER: 09/356931
;  PRIOR FILING DATE: 1990-07-19
;  PRIOR APPLICATION NUMBER: 08/920162
;  PRIOR FILING DATE: 1997-08-27
;  PRIOR APPLICATION NUMBER: 08/548998
;  PRIOR FILING DATE: 1995-10-27
;  PRIOR APPLICATION NUMBER: 08/616081
;  PRIOR FILING DATE: 1996-03-14
;  PRIOR APPLICATION NUMBER: 08/703675
;  PRIOR FILING DATE: 1996-08-27
;  PRIOR APPLICATION NUMBER: 08/897342
;  PRIOR FILING DATE: 1997-07-21
;  NUMBER OF SEQ ID NOS: 34
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 30
;   LENGTH: 5
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: peptide
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(5)
;   OTHER INFORMATION: Modified site, D amino acid
US-09-895-443A-30

  Query Match             8.6%;  Score 5;  DB 4;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 4
US-09-627-851B-47
; Sequence 47, Application US/09627851B
; Patent No. 6545131
; GENERAL INFORMATION:
```

```
;  APPLICANT: ISSACS, JOHN T.
;  APPLICANT:  DENMEADE, SAMUEL R.
;  APPLICANT:  CHRISTENSEN, S. BROGGER
;  APPLICANT:  LILJA, HANS
;  TITLE OF INVENTION: TISSUE SPECIFIC PRODRUG
;  FILE REFERENCE: 57111 (71699)
;  CURRENT APPLICATION NUMBER: US/09/627,851B
;  CURRENT FILING DATE:  2000-07-28
;  PRIOR APPLICATION NUMBER: 09/588,822
;  PRIOR FILING DATE: 2000-06-07
;  PRIOR APPLICATION NUMBER: 09/081,707
;  PRIOR FILING DATE: 1998-05-19
;  PRIOR APPLICATION NUMBER: 60/047,070
;  PRIOR FILING DATE: 1997-05-19
;  PRIOR APPLICATION NUMBER: 60/080,046
;  PRIOR FILING DATE: 1998-03-30
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-627-851B-47

  Query Match             8.6%;  Score 5;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          42 GKKRR 46
               |||||
Db           2 GKKRR 6


RESULT 5
US-08-621-803-221
; Sequence 221, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  221:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.382:
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-221

  Query Match             8.6%;  Score 5;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          13 LQLLK 17
               |||||
Db           4 LQLLK 8


RESULT 6
US-08-621-259A-213
; Sequence 213, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  213:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.382:
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-213

  Query Match             8.6%;  Score 5;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          13 LQLLK 17
               |||||
Db           4 LQLLK 8


RESULT 7
US-08-159-339A-343
; Sequence 343, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  343:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-343

  Query Match             8.6%;  Score 5;  DB 3;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          34 SLMLL 38
               |||||
Db           1 SLMLL 5


RESULT 8
US-08-159-339A-405
; Sequence 405, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  405:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-405

  Query Match             8.6%;  Score 5;  DB 3;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          34 SLMLL 38
               |||||
Db           3 SLMLL 7


RESULT 9
US-09-217-352-221
; Sequence 221, Application US/09217352
; Patent No. 6274344
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/217,352
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/621,803
;       FILING DATE:  22-MAR-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Borun, Michael F.
;       REGISTRATION NUMBER:  25,447
;       REFERENCE/DOCKET NUMBER:  27129/33199
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/474-6300
;       TELEFAX:  312/474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  221:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  9 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.382:
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-09-217-352-221

  Query Match             8.6%;  Score 5;  DB 3;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 LQLLK 17
              |||||
Db          4 LQLLK 8


RESULT 10
US-09-677-664B-213
; Sequence 213, Application US/09677664B
; Patent No. 6664231
;   GENERAL INFORMATION:
;       APPLICANT: Little II, Roger G
;                  Lim, Edward
;                  Fadem, Mitchell B.
;       TITLE OF INVENTION: Anti-Fungal Peptides
;       NUMBER OF SEQUENCES: 257
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McAndrews, Held & Malloy, Ltd.
;            STREET: 500 West Madison Street
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: United States of America
;            ZIP: 60661
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/677,664B
;            FILING DATE: 07-Mar-2003
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/227,659
;            FILING DATE: 08-Jan-1999
;       ATTORNEY/AGENT INFORMATION:
;            NAME: McNicholas, Janet M.
;            REGISTRATION NUMBER: 32,918
;            REFERENCE/DOCKET NUMBER: 11021US06
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312/775-8000
;            TELEFAX: 312/775-8100
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 213:
;     SEQUENCE CHARACTERISTICS:
;            LENGTH: 9 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: peptide
;       FEATURE:
;            NAME/KEY:  misc_feature
;            OTHER INFORMATION:  "XMP.382:
;       FEATURE:
;            NAME/KEY:  Modified-site
;            LOCATION:  C-Terminus
;            OTHER INFORMATION:  /label= Amidation
;            /note= "The C-Terminus is Amidated."
;       SEQUENCE DESCRIPTION: SEQ ID NO: 213:
US-09-677-664B-213

  Query Match             8.6%;  Score 5;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 LQLLK 17
              |||||
Db          4 LQLLK 8


RESULT 11
US-09-792-480-18
; Sequence 18, Application US/09792480
; Patent No. 6669951
; GENERAL INFORMATION:
;  APPLICANT: Rothbard, Jonathan B.
;  APPLICANT:  Wender, Paul A.
;  APPLICANT:  McGrane, P. Leo
;  APPLICANT:  Sista, Lalitha V.S.
;  APPLICANT:  Kirschberg, Thorsten A.
;  APPLICANT:  CellGate, Inc.
;  TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
;  TITLE OF INVENTION:  Across and Into Epithelial Tissues
;  FILE REFERENCE: 019801-000230US
;  CURRENT APPLICATION NUMBER: US/09/792,480
;  CURRENT FILING DATE:  2001-02-23
;  PRIOR APPLICATION NUMBER: US 09/648,400
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/150,510
;  PRIOR FILING DATE: 1999-08-24
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:A-54
;   OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
;   OTHER INFORMATION: region Tat-49-57
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: Xaa = fluorescein linked to amino group of
;   OTHER INFORMATION: aminohexanoic acid (Fl-ahx) attached to the
;   OTHER INFORMATION: N-terminal amino group of Arg
US-09-792-480-18

  Query Match             8.6%;  Score 5;  DB 4;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRRA 47
              |||||
```

```
Db          2 KKRRA 6


RESULT 12
US-08-248-839C-135
; Sequence 135, Application US/08248839C
; Patent No. 5843702
;  GENERAL INFORMATION:
;    APPLICANT:  McConnell, David
;    APPLICANT:  Devine, Kevin
;    APPLICANT:  O'Kane, Charles
;    TITLE OF INVENTION:  A Gene Expression System
;    NUMBER OF SEQUENCES:  185
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5843702o No. 5843702disk of No. 5843702th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10174-6401
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/248,839C
;      FILING DATE:  25-MAY-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gregg, Valeta A.
;      REGISTRATION NUMBER:  35,127
;      REFERENCE/DOCKET NUMBER:  3614.214-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  135:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Protein
US-08-248-839C-135

  Query Match             8.6%;  Score 5;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         34 SLMLL 38
              |||||
Db          3 SLMLL 7


RESULT 13
US-09-648-400A-18
; Sequence 18, Application US/09648400A
; Patent No. 6593292
; GENERAL INFORMATION:
;  APPLICANT: Rothbard, Jonathan B.
;  APPLICANT:  Wender, Paul A.
;  APPLICANT:  McGrane, P. Leo
;  APPLICANT:  Sista, Lalitha V.S.
;  APPLICANT:  Kirschberg, Thorsten A.
;  APPLICANT:  CellGate, Inc.
;  TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
;  TITLE OF INVENTION:  Across and Into Epithelial Tissues
;  FILE REFERENCE: 019801-000210US
;  CURRENT APPLICATION NUMBER: US/09/648,400A
;  CURRENT FILING DATE:  2000-08-24
;  PRIOR APPLICATION NUMBER: US 60/150,510
;  PRIOR FILING DATE: 1999-08-24
;  NUMBER OF SEQ ID NOS: 30
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:A-54
;   OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
;   OTHER INFORMATION: region Tat-49-57
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: Xaa = fluorescein linked to amino group of
;   OTHER INFORMATION: aminohexanoic acid (Fl-ahx)
US-09-648-400A-18

  Query Match             8.6%;  Score 5;  DB 4;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRRA 47
              |||||
Db          3 KKRRA 7


RESULT 14
US-09-839-329-4
; Sequence 4, Application US/09839329
; Patent No. 6664040
; GENERAL INFORMATION:
;  APPLICANT: Michael P. Sherman
;  APPLICANT:  Warner C. Greene
;  APPLICANT:  Carlos M.C. de No. 6664040ohna
;  APPLICANT:  Ulrich Schubert
;  APPLICANT:  Peter Henklein
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF
;  TITLE OF INVENTION:  A MOLECULE INTO A CELL
;  FILE REFERENCE: G&C 30448.91-US-U2
;  CURRENT APPLICATION NUMBER: US/09/839,329
;  CURRENT FILING DATE:  2001-04-20
;  PRIOR APPLICATION NUMBER: 60/206,610
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/267,827
;  PRIOR FILING DATE: 2001-02-09
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;  OTHER INFORMATION: Human immunodeficiency virus
US-09-839-329-4

  Query Match             8.6%;  Score 5;  DB 4;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         42 GKKRR 46
              |||||
Db          2 GKKRR 6


RESULT 15
US-10-209-421-18
; Sequence 18, Application US/10209421
; Patent No. 6759387
; GENERAL INFORMATION:
;  APPLICANT: Rothbard, Jonathan B.
;  APPLICANT:  Wender, Paul A.
;  APPLICANT:  McGrane, P. Leo
;  APPLICANT:  Sista, Lalitha V.S.
;  APPLICANT:  Kirschberg, Thorsten A.
```

```
;  APPLICANT:  CellGate, Inc.
;  TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
;  TITLE OF INVENTION:  Across and Into Epithelial Tissues
;  FILE REFERENCE: 019801-000211US
;  CURRENT APPLICATION NUMBER: US/10/209,421
;  CURRENT FILING DATE:  2002-07-30
;  PRIOR APPLICATION NUMBER: US 60/150,510
;  PRIOR FILING DATE: 1999-08-24
;  PRIOR APPLICATION NUMBER: US 09/648,400
;  PRIOR FILING DATE: 2000-08-24
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:A-54
;   OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
;   OTHER INFORMATION: region Tat-49-57
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (1)
;   OTHER INFORMATION: Xaa = fluorescein linked to amino group of
;   OTHER INFORMATION: aminohexanoic acid (Fl-ahx)
US-10-209-421-18

  Query Match             8.6%;  Score 5;  DB 4;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRRA 47
              |||||
Db          3 KKRRA 7


Search completed: October 16, 2005, 06:01:38
Job time : 56 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:45:39 ; Search time 118 Seconds
(without alignments)
251.700 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2548

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	4	6.9	4	1	FLRN_ANTEL	P58707	anthopleura
2	4	6.9	7	2	P92210	P92210	agropyron c
3	4	6.9	7	2	P92214	P92214	amblyopyrum
4	4	6.9	7	2	P92218	P92218	australopyr
5	4	6.9	7	2	P92221	P92221	bromus iner
6	4	6.9	7	2	P92226	P92226	crithopsis
7	4	6.9	7	2	P92372	P92372	haynaldia v
8	4	6.9	7	2	P92381	P92381	hordeum bra
9	4	6.9	7	2	P92385	P92385	hordeum mar
10	4	6.9	7	2	P92387	P92387	henrardia p
11	4	6.9	7	2	P92390	P92390	heteranthel
12	4	6.9	7	2	P92393	P92393	hordeum vul
13	4	6.9	7	2	P92403	P92403	lophopyrum
14	4	6.9	7	2	P92421	P92421	psathyrosta
15	4	6.9	7	2	P92425	P92425	pseudoroegn
16	4	6.9	7	2	P92427	P92427	peridictyon
17	4	6.9	7	2	P92430	P92430	aegilops ta
18	4	6.9	7	2	P92440	P92440	thinopyrum
19	4	6.9	7	2	P92442	P92442	taeniatheru
20	4	6.9	8	2	O15896	O15896	babesia bov
21	4	6.9	9	2	O88889	O88889	mus musculu
22	4	6.9	10	1	FARP_LOCMI	P38553	locusta mig
23	4	6.9	10	1	LCMS_LEUMA	P21144	leucophaea
24	4	6.9	10	1	NEMS_SARBU	P61850	sarcophaga
25	4	6.9	10	2	Q6A3T0	Q6a3t0	archangiopt
26	4	6.9	10	2	Q76ML8	Q76ml8	eurypharynx
27	3	5.2	4	1	FLRF_HIRME	P42561	hirudo medi
28	3	5.2	5	1	FARP_ARTTR	P41853	artioposthi
29	3	5.2	5	1	PAP2_PARMA	P81864	pardachirus
30	3	5.2	7	1	FAR1_HELTI	P41871	helisoma tr
31	3	5.2	7	1	FAR1_MACRS	P83274	macrobrachi
32	3	5.2	7	1	FAR1_PROCL	P38499	procambarus
33	3	5.2	7	1	FAR2_PROCL	P38498	procambarus
34	3	5.2	7	2	Q9YIQ9	Q9yiq9	human adeno
35	3	5.2	7	2	Q9YIR0	Q9yir0	human adeno
36	3	5.2	7	2	Q9YVE3	Q9yve3	human adeno
37	3	5.2	8	1	AKHG_GRYBI	P67785	gryllus bim
38	3	5.2	8	1	AKH_ROMMI	P67786	romalea mic
39	3	5.2	8	1	CAD1_ENTFA	P13268	enterococcu
40	3	5.2	8	1	FAR1_PANRE	P41872	panagrellus
41	3	5.2	8	1	FAR1_PENMO	P83316	penaeus mon
42	3	5.2	8	1	FAR2_MACRS	P83275	macrobrachi
43	3	5.2	8	1	FAR3_HOMAM	P41486	homarus ame
44	3	5.2	8	1	FAR4_HOMAM	P41487	homarus ame
45	3	5.2	8	1	HTF1_PERAM	P04548	periplaneta

ALIGNMENTS

RESULT 1
FLRN_ANTEL
ID FLRN_ANTEL STANDARD; PRT; 4 AA.
AC P58707;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Antho-RNamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC Nynantheae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RX MEDLINE=90319122; PubMed=1973541;
RA Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
RA Reinscheid R.K., Nothacker H.-P., Staley A.L.;
RT "Isolation of L-3-phenyllactyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
RT anemone neuropeptide containing an unusual amino-terminal blocking
RT group.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Neuron specific.
CC -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB; RANGE=1-4; NOTE=Ref.1.
DR PIR; A35779; A35779.
KW Amidation; Direct protein sequencing; Neuropeptide.
FT MOD_RES 1 1 3-phenyllactic acid.
FT MOD_RES 4 4 Asparagine amide.
SQ SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;

Query Match 6.9%; Score 4; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 FLRN 32
||||
Db 1 FLRN 4

RESULT 2
P92210
ID P92210 PRELIMINARY; PRT; 7 AA.
AC P92210;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Ribosomal protein 11 (Fragment).
GN Name=rps11;
OS Agropyron cristatum (Crested wheatgrass).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Agropyron.
OX NCBI_TaxID=4593;

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77771; CAB01391.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 3
P92214
ID   P92214      PRELIMINARY;      PRT;     7 AA.
AC   P92214;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Amblyopyrum muticum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Amblyopyrum.
OX   NCBI_TaxID=4595;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77756; CAB01346.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 4
P92218
ID   P92218      PRELIMINARY;      PRT;     7 AA.
AC   P92218;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Australopyrum retrofractum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Australopyrum.
OX   NCBI_TaxID=4597;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77767; CAB01379.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 5
P92221
ID   P92221      PRELIMINARY;      PRT;     7 AA.
AC   P92221;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Bromus inermis (Smooth brome grass).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Bromeae; Bromus.
OX   NCBI_TaxID=15371;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77759; CAB01355.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 6
P92226
ID   P92226      PRELIMINARY;      PRT;     7 AA.
AC   P92226;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
```

```
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Crithopsis delileana.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Crithopsis.
OX   NCBI_TaxID=37674;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77751; CAB01331.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 7
P92372
ID   P92372      PRELIMINARY;      PRT;     7 AA.
AC   P92372;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Haynaldia villosa.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Haynaldia.
OX   NCBI_TaxID=40247;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77741; CAB01301.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 8
P92381
ID   P92381      PRELIMINARY;      PRT;     7 AA.
AC   P92381;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Hordeum brachyantherum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=52712;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77761; CAB01361.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 9
P92385
ID   P92385      PRELIMINARY;      PRT;     7 AA.
AC   P92385; P92383;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Hordeum marinum subsp. gussoneanum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=98114;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77763; CAB01367.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          43 KKRR 46
               ||||
Db           3 KKRR 6


RESULT 10
P92387
ID   P92387      PRELIMINARY;      PRT;     7 AA.
AC   P92387;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Henrardia persica.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Henrardia.
OX   NCBI_TaxID=37678;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77748; CAB01322.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          43 KKRR 46
               ||||
Db           3 KKRR 6


RESULT 11
P92390
ID   P92390      PRELIMINARY;      PRT;     7 AA.
AC   P92390;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Heteranthelium piliferum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Heteranthelium.
OX   NCBI_TaxID=37679;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77750; CAB01328.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          43 KKRR 46
               ||||
Db           3 KKRR 6


RESULT 12
P92393
ID   P92393      PRELIMINARY;      PRT;     7 AA.
AC   P92393;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Hordeum vulgare (Barley).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Hordeum.
OX   NCBI_TaxID=4513;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77764; CAB01370.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          43 KKRR 46
               ||||
Db           3 KKRR 6


RESULT 13
P92403
ID   P92403      PRELIMINARY;      PRT;     7 AA.
AC   P92403;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Lophopyrum elongatum (Tall wheatgrass) (Argopyrum elongatum).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Thinopyrum.
OX   NCBI_TaxID=4588;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77743; CAB01307.1; -.
```

```
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 14
P92421
ID   P92421      PRELIMINARY;      PRT;     7 AA.
AC   P92421; P92419;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Psathyrostachys fragilis.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Psathyrostachys.
OX   NCBI_TaxID=37729;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77753; CAB01337.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


RESULT 15
P92425
ID   P92425      PRELIMINARY;      PRT;     7 AA.
AC   P92425;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein 11 (Fragment).
GN   Name=rps11;
OS   Pseudoroegneria spicata (Bluebunch wheatgrass) (Agropyron spicatum).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC   Triticeae; Pseudoroegneria.
OX   NCBI_TaxID=4604;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaves;
RX   MEDLINE=97271648; PubMed=9126564; DOI=10.1006/mpev.1996.0389;
RA   Petersen G., Seberg O.;
RT   "Phylogenetic analysis of the Triticeae (Poaceae) based on rpoA
RT   sequence data.";
RL   Mol. Phylogenet. Evol. 7:217-230(1997).
DR   EMBL; Z77744; CAB01310.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  894 MW;  6734140333277700 CRC64;

  Query Match             6.9%;  Score 4;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 KKRR 46
              ||||
Db          3 KKRR 6


Search completed: October 16, 2005, 05:59:29
Job time : 119 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:51:50 ; Search time 51 Seconds
(without alignments)
109.423 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1102

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4	6.9	4	2	A35779	neuropeptide Antho
2	4	6.9	8	2	PC4131	hypothetical prote
3	4	6.9	10	2	A56633	neomyosuppressin -
4	4	6.9	10	2	A32543	cardioexcitatory n
5	3	5.2	5	2	PT0525	T-cell receptor be
6	3	5.2	5	2	PT0577	T-cell receptor be
7	3	5.2	5	2	PT0700	T-cell receptor be
8	3	5.2	5	2	PT0565	T-cell receptor be
9	3	5.2	6	2	A11490	pyruvate kinase (E
10	3	5.2	6	2	PT0518	T-cell receptor be
11	3	5.2	7	2	S78024	ribosomal protein
12	3	5.2	7	2	I48086	DNA topoisomerase
13	3	5.2	7	2	A61081	tryptophyllin, bas
14	3	5.2	8	2	A44960	neuropeptide Led-C
15	3	5.2	8	2	A61348	red pigment-concen
16	3	5.2	8	2	S08995	hypertrehalosemic
17	3	5.2	8	2	A49823	adipokinetic hormo
18	3	5.2	8	2	A28004	adipokinetic hormo
19	3	5.2	8	2	A43976	hypertrehalosemic
20	3	5.2	8	2	B43976	hypertrehalosemic
21	3	5.2	8	2	T10077	hypothetical prote
22	3	5.2	8	2	S19288	acylase - Kluyvera
23	3	5.2	8	2	S21288	lectin - potato (f
24	3	5.2	8	2	A05169	neuropeptide M-I -
25	3	5.2	8	2	JS0316	leucokinin VI - Ma
26	3	5.2	8	2	S71919	alcohol dehydrogen
27	3	5.2	9	2	A61102	parathyroid hormon
28	3	5.2	9	2	PT0326	Ig heavy chain CRD
29	3	5.2	9	2	S19523	orf AB protein - S
30	3	5.2	9	2	A31576	xylose isomerase (
31	3	5.2	9	2	PC7076	spectrin alpha cha
32	3	5.2	10	2	A60624	angiotensin I - Ja
33	3	5.2	10	2	A60410	beta-neoendorphin
34	3	5.2	10	2	S08997	hypertrehalosemic
35	3	5.2	10	2	A60421	hypertrehalosemic
36	3	5.2	10	2	S08998	hypertrehalosemic
37	3	5.2	10	2	A26381	hypertrehalosemic
38	3	5.2	10	2	JN0440	peptide-N4-(N-acet
39	3	5.2	10	2	PQ0753	beta-fructofuranos
40	3	5.2	10	2	A61622	vitellogenin, 190k
41	3	5.2	10	2	C39111	Ig heavy chain C r
42	3	5.2	10	2	S10785	enamelin, 22K - bo
43	3	5.2	10	2	S70722	65.4K GTP-binding
44	3	5.2	10	2	PT0084	protein QA600021 -
45	3	5.2	10	2	B59272	peptide-N4-(N-acet

ALIGNMENTS

RESULT 1
A35779
neuropeptide Antho-RNamide - sea anemone (Anthopleura elegantissima)
C;Species: Anthopleura elegantissima
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C;Accession: A35779
R;Grimmelikhuijzen, C.J.P.; Rinehart, K.L.; Jacob, E.; Graff, D.; Reinscheid, R.K.; Noth
Proc. Natl. Acad. Sci. U.S.A. 87, 5410-5414, 1990
A;Title: Isolation of L-3-phenyllactyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea anemone ne
A;Reference number: A35779; MUID:90319122; PMID:1973541
A;Accession: A35779
A;Molecule type: protein
A;Residues: 1-4 <GRI>
A;Cross-references: UNIPROT:P58707
C;Comment: The L-3-phenyllactyl blocking group probably arises from an amino-terminal ph
C;Keywords: amidated carboxyl end; neuropeptide; phenyllactylation
F;1/Modified site: L-3-phenyllactic acid (Phe) #status experimental
F;4/Modified site: amidated carboxyl end (Asn) #status experimental

Query Match 6.9%; Score 4; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 FLRN 32
||||
Db 1 FLRN 4

RESULT 2
PC4131
hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: PC4131
R;Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.
Gene 167, 87-91, 1995
A;Title: Sequencing and characterization of the downstream region of the genes encoding
y for biosynthesis of heme d1.
A;Reference number: JC4552; MUID:96144254; PMID:8566817
A;Accession: PC4131
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-8 <KAW>
A;Cross-references: UNIPROT:P95412; DDBJ:D50473; NID:g1217594
A;Note: this ORF is not annotated in GenBank entry PSENIRC, release 113.0

Query Match 6.9%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 SRRL 23
||||

```
Db          5 SRRL 8


RESULT 3
A56633
neomyosuppressin - flesh fly (Sarcophaga bullata)
N;Alternate names: Neb-MS
C;Species: Sarcophaga bullata
C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C;Accession: A56633
R;Fonagy, A.; Schoofs, L.; Proost, P.; Van Damme, J.; Bueds, H.; De Loof, A.
Comp. Biochem. Physiol. C 102, 239-245, 1992
A;Title: Isolation, primary structure and synthesis of neomyosuppressin, a myoinhibiting
A;Reference number: A56633; MUID:93047886; PMID:1358537
A;Accession: A56633
A;Molecule type: protein
A;Residues: 1-10 <FON>
A;Cross-references: UNIPROT:P61850
A;Experimental source: head
A;Note: sequence extracted from NCBI backbone (NCBIP:119072)
C;Keywords: amidated carboxyl end; neuropeptide
F;10/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             6.9%;  Score 4;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 VFLR 31
              ||||
Db          6 VFLR 9


RESULT 4
A32543
cardioexcitatory neuropeptide - desert locust
C;Species: Schistocerca gregaria (desert locust)
C;Date: 20-Dec-1989 #sequence_revision 20-Dec-1989 #text_change 09-Jul-2004
C;Accession: A32543
R;Robb, S.; Packman, L.C.; Evans, P.D.
Biochem. Biophys. Res. Commun. 160, 850-856, 1989
A;Title: Isolation, primary structure and bioactivity of SchistoFLRf-amide, a FMRF-amide
A;Reference number: A32543; MUID:89246543; PMID:2719702
A;Accession: A32543
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <ROB>
A;Cross-references: UNIPROT:P38553
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end; neuropeptide
F;10/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             6.9%;  Score 4;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         28 VFLR 31
              ||||
Db          6 VFLR 9


RESULT 5
PT0525
T-cell receptor beta chain V-D-J region (100-4J) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0525
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0525
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             5.2%;  Score 3;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         55 RGT 57
              |||
Db          3 RGT 5


RESULT 6
PT0577
T-cell receptor beta chain V-D-J region (141-1BC) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0577; PT0574
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0577
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1BC
A;Accession: PT0574
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FE2>
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1Q
C;Keywords: T-cell receptor

  Query Match             5.2%;  Score 3;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 SSR 21
              |||
Db          2 SSR 4


RESULT 7
PT0700
T-cell receptor beta chain V-D-J region (161-2A) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0700
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0700
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             5.2%;  Score 3;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 SSR 21
              |||
Db          2 SSR 4


RESULT 8
PT0565
T-cell receptor beta chain V-D-J region (141-1CF) - mouse (fragment)
```

```
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: PT0565
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0565
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPROT:Q8BZQ7
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             5.2%;  Score 3;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 SSR 21
              |||
Db          2 SSR 4


RESULT 9
A11490
pyruvate kinase (EC 2.7.1.40) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 03-Mar-1995
C;Accession: A11490
R;Hjelmquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.
Biochem. Biophys. Res. Commun. 61, 559-563, 1974
A;Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase ph
A;Reference number: A11490; MUID:75127438; PMID:4375989
A;Accession: A11490
A;Molecule type: protein
A;Residues: 1-6 <HJE>
A;Experimental source: liver
C;Keywords: glycolysis; phosphotransferase

  Query Match             5.2%;  Score 3;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 RRA 47
              |||
Db          2 RRA 4


RESULT 10
PT0518
T-cell receptor beta chain V-D-J region (100-4B) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0518
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0518
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-6 <FEE>
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             5.2%;  Score 3;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         46 RAT 48
              |||
Db          3 RAT 5


RESULT 11
S78024
ribosomal protein YmL21, mitochondrial - yeast (Saccharomyces cerevisiae) (fragment)
C;Species: Saccharomyces cerevisiae
C;Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 14-Nov-1997
C;Accession: S78024
R;Kitakawa, M.; Graack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wittma
Eur. J. Biochem. 245, 449-456, 1997
A;Title: Identification and characterization of the genes for mitochondrial ribosomal pi
A;Reference number: S78018; MUID:97296414; PMID:9151978
A;Accession: S78024
A;Molecule type: protein
A;Residues: 1-7 <KIT>
C;Genetics:
A;Genome: nuclear
C;Keywords: mitochondrion; protein biosynthesis; ribosome

  Query Match             5.2%;  Score 3;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         15 LLK 17
              |||
Db          5 LLK 7


RESULT 12
I48086
DNA topoisomerase II alpha - Chinese hamster (fragment)
C;Species: Cricetulus griseus (Chinese hamster)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I48086
R;Ng, S.W.; Eder, J.P.; Schnipper, L.E.; Chan, V.T.W.
J. Biol. Chem. 270, 25850-25858, 1995
A;Title: Molecular cloning and characterization of the promoter for the Chinese hamster
A;Reference number: I48086; MUID:96029684; PMID:7592770
A;Accession: I48086
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-7 <RES>
A;Cross-references: EMBL:U34196; NID:g1041231; PIDN:AAC52315.1; PID:g1041232

  Query Match             5.2%;  Score 3;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 SPL 9
              |||
Db          4 SPL 6


RESULT 13
A61081
tryptophyllin, basic - Rohde's leaf frog
C;Species: Phyllomedusa rohdei (Rohde's leaf frog)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Aug-2000
C;Accession: A61081
R;Montecucchi, P.C.; Vincenti, M.; Lazzarini, A.M.; Rusconi, L.; Erspamer, V.
Int. J. Pept. Protein Res. 33, 391-395, 1989
A;Title: Isolation, structure determination and synthesis of a novel tryptophan-containi
A;Reference number: A61081
A;Accession: A61081
A;Molecule type: protein
A;Residues: 1-7 <MON>
C;Comment: The biological activity of this peptide was not determined.
C;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end; hydroxyproline; skin
F;3/Modified site: 4-hydroxyproline (Pro) #status experimental
F;7/Modified site: amidated carboxyl end (Pro) #status experimental

  Query Match             5.2%;  Score 3;  DB 2;  Length 7;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 PPS 7
              |||
Db          2 PPS 4


RESULT 14
A44960
neuropeptide Led-CC-I - Colorado potato beetle
C;Species: Leptinotarsa decemlineata (Colorado potato beetle)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C;Accession: A44960
R;Gaede, G.; Kellner, R.
Peptides 10, 1287-1289, 1989
A;Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and
A;Reference number: A44960; MUID:90160053; PMID:2576128
A;Accession: A44960
A;Molecule type: protein
A;Residues: 1-8 <GAE>
A;Cross-references: UNIPROT:P04548
C;Superfamily: adipokinetic hormone
C;Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

  Query Match             5.2%;  Score 3;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 NFS 34
              |||
Db          3 NFS 5


RESULT 15
A61348
red pigment-concentrating hormone - northern shrimp
N;Alternate names: blanching hormone
C;Species: Pandalus borealis (northern shrimp)
C;Date: 02-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 09-Jul-2004
C;Accession: A61348; S07139
R;Fernlund, P.; Josefsson, L.
Science 177, 173-175, 1972
A;Title: Crustacean color-change hormone: amino acid sequence and chemical synthesis.
A;Reference number: A61348; MUID:72228738; PMID:5041363
A;Accession: A61348
A;Molecule type: protein
A;Residues: 1-8 <FER1>
A;Cross-references: UNIPROT:P08939
R;Fernlund, P.
Biochim. Biophys. Acta 371, 304-311, 1974
A;Title: Structure of the red-pigment-concentrating hormone of the shrimp, Pandalus bore
A;Reference number: S07139; MUID:75054965; PMID:4433569
A;Accession: S07139
A;Molecule type: protein
A;Residues: 'E',2-8 <FER2>
A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C;Comment: This neuropeptide, isolated from the eyestalks of the shrimp and active in pi
zed pigment-containing cells.
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; blocked amino end; hormone; neuropeptide; pyroglutami
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental

  Query Match             5.2%;  Score 3;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         32 NFS 34
              |||
Db          3 NFS 5
```

Search completed: October 16, 2005, 06:00:31
Job time : 52 secs

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:44:34 ; Search time 117 Seconds
(without alignments)
191.727 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 58
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 465227

Minimum DB seq length: 0
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6	10.3	9	8	ADH89809	Adh89809 Cell pene
2	5	8.6	5	2	AAW51344	Aaw51344 Peptide #
3	5	8.6	5	2	AAW51338	Aaw51338 Peptide #
4	5	8.6	5	3	AAY49996	Aay49996 Natural b
5	5	8.6	5	4	AAE12536	Aae12536 APP70 abe
6	5	8.6	5	4	AAB80747	Aab80747 hK2 cleav
7	5	8.6	5	6	ABG73485	Abg73485 Natural b
8	5	8.6	7	4	AAB80706	Aab80706 Human gla
9	5	8.6	7	4	AAB80669	Aab80669 Human gla
10	5	8.6	7	5	AAU85856	Aau85856 Prostate-
11	5	8.6	7	7	ADB79449	Adb79449 Parapoxvi
12	5	8.6	7	8	ADN08587	Adn08587 Ceramic t
13	5	8.6	7	8	ADP74892	Adp74892 Parapoxvi
14	5	8.6	8	3	AAB06282	Aab06282 proSKI-1
15	5	8.6	9	2	AAR59128	Aar59128 Peptide f
16	5	8.6	9	2	AAY38039	Aay38039 Hepatitis
17	5	8.6	9	2	AAY38106	Aay38106 Hepatitis
18	5	8.6	9	2	AAW44612	Aaw44612 Anti-fung
19	5	8.6	9	2	AAW43777	Aaw43777 Bacterici
20	5	8.6	9	2	AAY45608	Aay45608 Immunogen
21	5	8.6	9	2	AAY45674	Aay45674 Immunogen
22	5	8.6	9	2	AAY46768	Aay46768 Immunogen
23	5	8.6	9	2	AAY00589	Aay00589 Antifunga
24	5	8.6	9	4	AAB65513	Aab65513 Anti-fung
25	5	8.6	9	4	AAU00796	Aau00796 HIV-1 Tat
26	5	8.6	9	5	ABB08361	Abb08361 Synthetic
27	5	8.6	9	5	AAO15791	Aao15791 Human imm
28	5	8.6	9	5	ABP54737	Abp54737 HIV-1 tat
29	5	8.6	9	6	AAE35004	Aae35004 Mouse imm
30	5	8.6	9	7	ABR82213	Abr82213 Human ant
31	5	8.6	9	7	ADK65211	Adk65211 Human PIG
32	5	8.6	9	8	ADE97590	Ade97590 Immunogen
33	5	8.6	9	8	ADH89810	Adh89810 Cell pene
34	5	8.6	9	8	ADH44385	Adh44385 INK4a imm
35	5	8.6	9	8	ADH44376	Adh44376 INK4a imm
36	5	8.6	9	8	ADK84035	Adk84035 Human 191
37	5	8.6	9	8	ADK84519	Adk84519 Human 191
38	5	8.6	9	8	ADK85026	Adk85026 Human 191
39	5	8.6	9	8	ADK85074	Adk85074 Human 191
40	5	8.6	9	8	ADK87550	Adk87550 Human 191
41	5	8.6	9	8	ADK88447	Adk88447 Human 191
42	5	8.6	9	8	ADK83950	Adk83950 Human 191
43	5	8.6	9	8	ADK88445	Adk88445 Human 191
44	5	8.6	9	8	ADK84560	Adk84560 Human 191
45	5	8.6	9	8	ADK85620	Adk85620 Human 191

ALIGNMENTS

RESULT 1
ADH89809
ID ADH89809 standard; peptide; 9 AA.
XX
AC ADH89809;
XX
DT 22-APR-2004 (first entry)
XX
DE Cell penetrating peptide (CPP) identification method-related peptide 111.
XX
KW cell-penetrating peptide; CPP; bulk property value Z-E; Z-E1; Z-E2; Z-E3;
KW Z-E4; Z-E5; antidiabetic; neuroprotective; nootropic; antiparkinsonian;
KW cardiant; cytostatic; tranquiliser; immunosuppressive; antidepressant;
KW anticonvulsant; antiinflammatory; analgesic; neuroleptic;
KW ophthalmological; antiulcer; cell-penetration; infectious disease;
KW diabetes type I; diabetes type II; Alzheimer's disease;
KW Parkinson's disease; cancer; prion disease; cardiovascular disease;
KW signal transduction.
XX
OS Unidentified.
XX
PN WO2003106491-A2.
XX
PD 24-DEC-2003.
XX
PF 18-JUN-2003; 2003WO-IB003163.
XX
PR 18-JUN-2002; 2002SE-00001863.
PR 25-JUN-2002; 2002US-0391788P.
XX
PA (CEPE-) CEPEP AB.
XX
PI Haellbrink M, Pooga M, Metsis M, Kogerman P, Valkna A, Meikas A;
PI Lindgren M, Graeslund A, Eriksson G, Oestensson CG, Budihna M;
PI Zorko M, Elmquist A, Soomets U, Lundberg P, Jaerver P, Saar K;
PI El-Andaloussi S, Kilk K, Langel U;
XX
DR WPI; 2004-090832/09.
XX
PT Predicting, designing, detecting, and/or verifying novel cell-penetrating
PT peptide based on assessment of bulk property value of sequences of cell-
PT penetrating peptide.
XX
PS Disclosure; Page 31; 148pp; English.
XX
CC This invention relates to a novel method of identifying, designing,
CC detecting, and/or verifying novel cell-penetrating peptide (CPP) based on
CC assessment of bulk property value Z-E of sequences of CPP comprising 5 or

```
CC   more individual average interval values Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5,
CC   where Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5 are average values of the
CC   respective descriptor values for the residues in the amino acid sequence.
CC   The invention may be useful for the development of compounds with an
CC   antidiabetic, neuroprotective, nootropic, antiparkinsonian, cardiant,
CC   cytostatic, tranquiliser, immunosuppressive, antidepressant,
CC   anticonvulsant, antiinflammatory, analgesic, neuroleptic,
CC   ophthalmological or antiulcer activity as a stimulator of cell-
CC   penetration. The method of the invention is useful for identifying a cell
CC   -penetrating peptide or protein and/or a cell-penetrating fragment of a
CC   peptide or protein. In addition, the invention may be useful for checking
CC   cellular penetration properties of a peptide, for producing a cell-
CC   penetrating and functional protein-mimicking peptide and for de novo
CC   design and production of an artificial cell-penetrating and/or and
CC   artificial cell-penetrating and functional protein-mimicking peptide.
CC   Compositions developed within the scope of the present invention may be
CC   useful for treating infectious diseases, diabetes type I, diabetes type
CC   II, Alzheimer's disease, Parkinson's diseases, cancer, prion disease,
CC   cardiovascular disease or disorders resulting from perturbed signal
CC   transduction. The method of the invention is fast, efficient and reliable
CC   for identifying, detecting, designing CPPs and for screening cellular
CC   uptake of a broad variety of CPPs in vitro and in vivo. The present
CC   sequence is that of a peptide which is related to the invention.
XX
SQ   Sequence 9 AA;

  Query Match             10.3%;  Score 6;  DB 8;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 FLQLLK 17
              ||||||
Db          4 FLQLLK 9


RESULT 2
AAW51344
ID   AAW51344 standard; peptide; 5 AA.
XX
AC   AAW51344;
XX
DT   14-AUG-1998  (first entry)
XX
DE   Peptide #27 useful as modulator of beta-amyloid peptide aggregation.
XX
KW   Natural beta-amyloid peptide; aggregation; D-amino acid;
KW   Alzheimer's disease; beta-amyloidosis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "D-form residue, N-terminal acetyl"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 3
FT                   /note= "D-form residue"
FT   Misc-difference 4
FT                   /note= "D-form residue"
FT   Modified-site   5
FT                   /note= "D-form residue, C-terminal amide"
XX
PN   WO9808868-A1.
XX
PD   05-MAR-1998.
XX
PF   27-AUG-1997;   97WO-US015166.
XX
PR   27-AUG-1996;   96US-00703675.
PR   21-JUL-1997;   97US-00897342.
XX
PA   (PRAE-) PRAECIS PHARM INC.
XX
PI   Findeis MA,  Gefter ML,  Musso G,  Signer ER,  Wakefield J;
PI   Molineaux S,  Chin J,  Lee J,  Kelley M,  Komar-Panicucci S;
PI   Arico-Muendel CC,  Phillips K,  Hayward NJ;
XX
DR   WPI; 1998-216936/19.
XX
PT   Peptide compounds which are preferably based on beta-amyloid peptide(s) -
PT   are useful in treatment of disorders related to beta-amyloidosis,
PT   especially Alzheimer's disease.
XX
PS   Claim 8; Page 80; 92pp; English.
XX
CC   The invention relates to peptides that modulate natural beta-amyloid
CC   peptide aggregation. The modulators of the invention comprise a peptide
CC   preferably based on a beta-amyloid peptide, that is comprised entirely of
CC   D-amino acids. Preferably the peptide comprises 3-5 D-amino acid residues
CC   and includes at least two D-amino acid residues independently selected
CC   from the group consisting of D-leucine, D-phenylalanine and D-valine.
CC   Preferred amino-terminal modifying groups include cyclic, heterocylic,
CC   polycyclic and branched alkyl groups. Preferred carboxy-terminal
CC   modifying groups include an amide group, an alkyl amide group, an aryl
CC   amide group or a hydroxy group. The peptides may be used to treat
CC   disorders associated with beta-amyloidosis, especially Alzheimer's
CC   disease. They may also be used in methods for detecting the presence of
CC   beta-amyloid peptides in biological samples. The present sequence
CC   represents a specifically claimed peptide
XX
SQ   Sequence 5 AA;

  Query Match              8.6%;  Score 5;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 3
AAW51338
ID   AAW51338 standard; peptide; 5 AA.
XX
AC   AAW51338;
XX
DT   14-AUG-1998  (first entry)
XX
DE   Peptide #21 useful as modulator of beta-amyloid peptide aggregation.
XX
KW   Natural beta-amyloid peptide; aggregation; D-amino acid;
KW   Alzheimer's disease; beta-amyloidosis.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D-form residue"
FT   Misc-difference 2
FT                   /note= "D-form residue"
FT   Misc-difference 3
FT                   /note= "D-form residue"
FT   Misc-difference 4
FT                   /note= "D-form residue"
FT   Modified-site   5
FT                   /note= "D-form residue, C-terminal amide"
XX
PN   WO9808868-A1.
XX
PD   05-MAR-1998.
XX
PF   27-AUG-1997;   97WO-US015166.
XX
```

```
PR   27-AUG-1996;   96US-00703675.
PR   21-JUL-1997;   97US-00897342.
XX
PA   (PRAE-) PRAECIS PHARM INC.
XX
PI   Findeis MA,  Gefter ML,  Musso G,  Signer ER,  Wakefield J;
PI   Molineaux S,  Chin J,  Lee J,  Kelley M,  Komar-Panicucci S;
PI   Arico-Muendel CC,  Phillips K,  Hayward NJ;
XX
DR   WPI; 1998-216936/19.
XX
PT   Peptide compounds which are preferably based on beta-amyloid peptide(s) -
PT   are useful in treatment of disorders related to beta-amyloidosis,
PT   especially Alzheimer's disease.
XX
PS   Claim 8; Page 78; 92pp; English.
XX
CC   The invention relates to peptides that modulate natural beta-amyloid
CC   peptide aggregation. The modulators of the invention comprise a peptide
CC   preferably based on a beta-amyloid peptide, that is comprised entirely of
CC   D-amino acids. Preferably the peptide comprises 3-5 D-amino acid residues
CC   and includes at least two D-amino acid residues independently selected
CC   from the group consisting of D-leucine, D-phenylalanine and D-valine.
CC   Preferred amino-terminal modifying groups include cyclic, heterocylic,
CC   polycyclic and branched alkyl groups. Preferred carboxy-terminal
CC   modifying groups include an amide group, an alkyl amide group, an aryl
CC   amide group or a hydroxy group. The peptides may be used to treat
CC   disorders associated with beta-amyloidosis, especially Alzheimer's
CC   disease. They may also be used in methods for detecting the presence of
CC   beta-amyloid peptides in biological samples. The present sequence
CC   represents a specifically claimed peptide
XX
SQ   Sequence 5 AA;

  Query Match             8.6%;  Score 5;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 4
AAY49996
ID   AAY49996 standard; peptide; 5 AA.
XX
AC   AAY49996;
XX
DT   04-FEB-2000  (first entry)
XX
DE   Natural beta amyloid peptide aggregation modulating compound #3.
XX
KW   Beta amyloid peptide aggregation; modulation; D-amino acid; diagnosis;
KW   retro-inverso isomer; amyloidogenic disease; Alzheimer's disease;
KW   Down's syndrome; hereditary cerebral haemorrhage; HCHWA-D;
KW   amyloidosis-Dutch-type.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .5
FT                   /note= "D-form residues"
FT   Modified-site   5
FT                   /note= "amidated"
XX
PN   US5985242-A.
XX
PD   16-NOV-1999.
XX
PF   27-AUG-1997;   97US-00920162.
XX
PR   27-OCT-1995;   95US-00548998.
PR   14-MAR-1996;   96US-00616081.
PR   27-AUG-1996;   96US-00703675.
PR   21-JUL-1997;   97US-00897342.
XX
PA   (PRAE-) PRAECIS PHARM INC.
XX
PI   Wakefield J,  Molineaux S,  Signer ER,  Kelley M,  Komar-Panicucci S;
PI   Musso G,  Phillips K,  Hayward NJ,  Gefter ML,  Findeis MA,  Lee J;
PI   Arico-Muendel CC,  Chin J;
XX
DR   WPI; 2000-022266/02.
XX
PT   Compound comprising a peptidic structure, an amino-terminal modifying
PT   group and a carboxy-terminal modifying group, useful for treating
PT   Alzheimer's disease.
XX
PS   Claim 8; Col 63; 40pp; English.
XX
CC   The present invention describes a compound of formula A-(Xaa)-B,
CC   comprising a peptidic structure (Xaa), an amino-terminal modifying group
CC   (A); and a carboxy-terminal modifying group (B). AAY49973 to AAY49994
CC   represent specifically claimed examples of (Xaa). Also described is a
CC   method for inhibiting aggregation of natural beta-amyloid peptides and
CC   treating Alzheimer's disease, comprising contacting the amyloid peptides
CC   with A-(Xaa)-B; and a method for detecting the presence of natural beta-
CC   amyloid peptides in a biological sample by contacting with A-(Xaa)-B. The
CC   compound is useful for treating Alzheimer's disease and for detecting the
CC   presence of natural beta-amyloid peptides in a biological sample. The
CC   compound can also be used prophylactically or therapeutically to treat
CC   other clinical occurrences of beta-amyloid deposition, such as in Down's
CC   syndrome individuals and in patients with hereditary cerebral haemorrhage
CC   with amyloidosis-Dutch-type (HCHWA-D). The present sequence represents an
CC   example of the compound from the present invention
XX
SQ   Sequence 5 AA;

  Query Match             8.6%;  Score 5;  DB 3;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 5
AAE12536
ID   AAE12536 standard; peptide; 5 AA.
XX
AC   AAE12536;
XX
DT   03-JAN-2002  (first entry)
XX
DE   APP70 abeta peptide (residues 17-21) inverso isomer mutant (V18F; A21L).
XX
KW   Beta-amyloid peptide; beta-AP; Abeta peptide; therapy; beta-amyloidosis;
KW   Alzheimer's disease; neuroprotective; beta-amyloid precursor protein;
KW   APP-770; nootropic; mutant; mutein.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .5
FT                   /note= "D-form residues"
FT   Modified-site   1
FT                   /note= "Optionally modified with 4-hydroxy benzoyl or
FT                   acetyl"
FT   Misc-difference 2
FT                   /note= "Wild-type Val substituted with Phe; corresponds
FT                   to position 18 of natural Abeta peptide"
FT   Misc-difference 5
FT                   /note= "Wild-type Ala substituted with Leu; corresponds
```

FT to position 21 of natural Abeta peptide; Optionally C-
FT terminal amide"
XX
PN US6277826-B1.
XX
PD 21-AUG-2001.
XX
PF 19-JUL-1999; 99US-00356931.
XX
PR 27-AUG-1996; 96US-00703675.
PR 27-AUG-1997; 97US-00920162.
XX
PA (PRAE-) PRAECIS PHARM INC.
XX
PI Findeis MA, Gefter ML, Musso G, Signer ER, Wakefield J;
PI Molineaux S, Chin J, Lee J, Kelley M, Komar-Panicucci S;
PI Arico-Muendel CC, Phillips K, Hayward NJ;
XX
DR WPI; 2001-637856/73.
XX
PT Modulator compound for treating disorders associated with beta-
PT amyloidosis e.g. Alzheimer's disease, comprises a beta-amyloid peptide
PT containing D-amino acids.
XX
PS Example 11; Col 48; 41pp; English.
XX
CC The patent discloses compounds and pharmaceutical compositions thereof,
CC that can bind to natural beta amyloid peptide (beta-AP; Abeta peptide),
CC modulate the aggregation of natural beta-AP and/or inhibit the cyto-
CC toxicity of natural beta-APs. The beta-amyloid modulator compounds of the
CC invention comprise a peptide, preferably based on beta-AP, that is
CC composed entirely of D-amino acids. The modulators of the invention are
CC useful for treating a disorder associated with beta-amyloidosis such as
CC Alzheimer's disease. The present sequence is an inverso isomer mutant
CC (V18F; A21L) of Abeta peptide. Beta AP is a cleavage product of beta
CC amyloid precursor protein (APP-770; residues 17-21)
XX
SQ Sequence 5 AA;

Query Match 8.6%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LFFFL 13
|||||
Db 1 LFFFL 5

RESULT 6
AAB80747
ID AAB80747 standard; peptide; 5 AA.
XX
AC AAB80747;
XX
DT 02-MAY-2001 (first entry)
XX
DE hK2 cleavage site peptide #38.
XX
KW Cleavage; kallikrein 2; hK2; prodrug.
XX
OS Synthetic.
XX
PN WO200109165-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US040496.
XX
PR 29-JUL-1999; 99US-0146316P.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Denmeade SR, Isaacs JT, Lilja H, Christensen SB;
XX
DR WPI; 2001-191450/19.
XX
PT New peptides containing cleavage sites specifically cleaved by human
PT kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS Claim 8; Page 30; 38pp; English.
XX
CC The present invention relates to a peptide comprising an amino acid
CC sequence having a cleavage site specific for an enzyme having a
CC proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC amino acids in length. The invention is useful for producing a prodrug
CC which involves linking a drug which contains a primary amine to the
CC peptide, in which the linking of the peptide to the drug inhibits the
CC therapeutic activity of the drug
XX
SQ Sequence 5 AA;

Query Match 8.6%; Score 5; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 GKKRR 46
|||||
Db 1 GKKRR 5

RESULT 7
ABG73485
ID ABG73485 standard; peptide; 5 AA.
XX
AC ABG73485;
XX
DT 10-MAY-2003 (first entry)
XX
DE Natural beta-amyloid peptide modulator compound #28.
XX
KW Natural beta-amyloid peptide; aggregation inhibition; beta-amyloidosis;
KW Alzheimer's disease; beta-amyloid deposition; Down's syndrome;
KW hereditary cerebral haemorrhage with amyloidosis-Dutch-type; HCHWA-D;
KW neurological impairment; neuroprotective; nootropic.
XX
OS Synthetic.
XX
PN US2002103134-A1.
XX
PD 01-AUG-2002.
XX
PF 29-JUN-2001; 2001US-00895443.
XX
PR 27-OCT-1995; 95US-00548998.
PR 14-MAR-1996; 96US-00616081.
PR 27-AUG-1996; 96US-00703675.
PR 21-JUL-1997; 97US-00897342.
PR 27-AUG-1997; 97US-00920162.
PR 19-JUL-1999; 99US-00356931.
XX
PA (PRAE-) PRAECIS PHARM INC.
XX
PI Findeis MA, Gefter ML, Musso G, Signer ER, Wakefield J;
PI Molineaux S, Chin J, Lee J, Kelley M, Komar-Panicucci S;
PI Arico-Muendel CC, Phillips K, Hayward NJ;
XX
DR WPI; 2003-298609/29.
XX
PT Novel compound that modulate natural beta-amyloid peptide aggregation, is
PT useful for treating amyloidogenic diseases e.g. Alzheimer's disease.
XX
PS Claim 8; Page 36; 42pp; English.
XX
CC The invention relates to a compound that modulates natural beta-amyloid
CC peptide aggregation. The peptide is useful for inhibiting aggregation of

```
CC   natural beta-amyloid peptides and for treating a disorder associated with
CC   beta-amyloidosis, e.g. Alzheimer's disease. The compound is also useful
CC   for detecting the presence or absence of natural beta-amyloid peptides in
CC   a biological sample, by contacting the biological sample with the
CC   compound in vitro, where the compound is labelled with a detectable
CC   substance and detecting the compound bound to natural beta-amyloid
CC   peptides to thus detect the presence or absence of natural beta-amyloid
CC   peptides. The compound is preferably labelled with radioactive technetium
CC   or radioactive iodine and is administered to the subject. The compound is
CC   also useful for prophylactically or therapeutically treating other
CC   clinical occurrences of beta-amyloid deposition such as in Down's
CC   syndrome individuals and in patients with hereditary cerebral haemorrhage
CC   with amyloidosis-Dutch-type (HCHWA-D), and for therapeutically allowing
CC   for the sequestration of beta-amyloid proteins at sites that do not lead
CC   to neurological impairment. This sequence represents a natural beta-
CC   amyloid peptide modulator compound of the invention
XX
SQ   Sequence 5 AA;

  Query Match             8.6%;  Score 5;  DB 6;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 LFFFL 13
              |||||
Db          1 LFFFL 5


RESULT 8
AAB80706
ID   AAB80706 standard; peptide; 7 AA.
XX
AC   AAB80706;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human glandular kallikrein 2 substrate peptide #36.
XX
KW   Cleavage; kallikrein 2; hK2; prodrug.
XX
OS   Synthetic.
XX
PN   WO200109165-A2.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-US040496.
XX
PR   29-JUL-1999;   99US-0146316P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Denmeade SR,  Isaacs JT,  Lilja H,  Christensen SB;
XX
DR   WPI; 2001-191450/19.
XX
PT   New peptides containing cleavage sites specifically cleaved by human
PT   kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT   cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS   Example 8; Page 29; 38pp; English.
XX
CC   The present invention relates to a peptide comprising an amino acid
CC   sequence having a cleavage site specific for an enzyme having a
CC   proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC   amino acids in length. The invention is useful for producing a prodrug
CC   which involves linking a drug which contains a primary amine to the
CC   peptide, in which the linking of the peptide to the drug inhibits the
CC   therapeutic activity of the drug
XX
SQ   Sequence 7 AA;

  Query Match             8.6%;  Score 5;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         42 GKKRR 46
              |||||
Db          2 GKKRR 6


RESULT 9
AAB80669
ID   AAB80669 standard; peptide; 7 AA.
XX
AC   AAB80669;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Human glandular kallikrein 2 cleavage site peptide #30.
XX
KW   Cleavage; kallikrein 2; hK2; prodrug.
XX
OS   Synthetic.
XX
PN   WO200109165-A2.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-US040496.
XX
PR   29-JUL-1999;   99US-0146316P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Denmeade SR,  Isaacs JT,  Lilja H,  Christensen SB;
XX
DR   WPI; 2001-191450/19.
XX
PT   New peptides containing cleavage sites specifically cleaved by human
PT   kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT   cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS   Disclosure; Page 8; 38pp; English.
XX
CC   The present invention relates to a peptide comprising an amino acid
CC   sequence having a cleavage site specific for an enzyme having a
CC   proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC   amino acids in length. The invention is useful for producing a prodrug
CC   which involves linking a drug which contains a primary amine to the
CC   peptide, in which the linking of the peptide to the drug inhibits the
CC   therapeutic activity of the drug
XX
SQ   Sequence 7 AA;

  Query Match             8.6%;  Score 5;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         42 GKKRR 46
              |||||
Db          2 GKKRR 6


RESULT 10
AAU85856
ID   AAU85856 standard; peptide; 7 AA.
XX
AC   AAU85856;
XX
DT   08-MAY-2002  (first entry)
XX
DE   Prostate-specific antigen (PSA)-specific peptide #42.
XX
KW   Prostate-specific antigen; PSA; cytostatic; sesquiterpene-gamma-lactone;
KW   cell proliferative disorder; malignant cell; non-malignant cell;
```

KW prostate cancer; thapsigargin.
XX
OS Synthetic.
XX
PN WO200193861-A1.
XX
PD 13-DEC-2001.
XX
PF 28-JUL-2000; 2000WO-US020542.
XX
PR 07-JUN-2000; 2000US-00588822.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Isaacs JT, Lilja H, Denmeade SR, Christensen SB;
XX
DR WPI; 2002-216722/27.
XX
PT New sesquiterpene-gamma-lactone comprising thapsigargin, is useful for
PT treating cell proliferative disorders.
XX
PS Disclosure; Page 18; 72pp; English.
XX
CC The invention relates to a novel therapeutic sesquiterpene-gamma-lactone
CC comprising thapsigargin, and prostate antigen-specific (PSA) peptides.
CC These are used for treating cell proliferative disorders including
CC malignant as well as non-malignant cell populations, such as in prostate
CC cancer. The derivative can be linked to a variety of carrier moieties,
CC such as the peptides of the invention. The linkage converts the
CC derivative into an inactive prodrug. AAU85818-AAU85861 represent the PSA-
CC specific peptides of the invention
XX
SQ Sequence 7 AA;

Query Match 8.6%; Score 5; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 GKKRR 46
|||||
Db 2 GKKRR 6

RESULT 11
ADB79449
ID ADB79449 standard; peptide; 7 AA.
XX
AC ADB79449;
XX
DT 04-DEC-2003 (first entry)
XX
DE Parapoxvirus ORF 12r N-terminal peptide.
XX
KW virucide; anti-HIV; hepatotropic; antiinflammatory; cytostatic;
KW vulnerary; antiasthmatic; antiallergic; dermatological; antidiabetic;
KW immunosuppressive; antirheumatic; antiarthritic; thyromimetic;
KW protozoacide; amoebicide; antibacterial; gene therapy; virus;
KW viral infections; non-viral infections; proliferative disease;
KW inflammatory disease; allergic disease; autoimmune disease.
XX
OS Parapoxvirus.
XX
PN WO2003006654-A2.
XX
PD 23-JAN-2003.
XX
PF 12-JUN-2002; 2002WO-EP006440.
XX
PR 13-JUN-2001; 2001NZ-00512341.
XX
PA (FARB) BAYER AG.
XX
PI Weber O, Friederichs SM, Siegling A, Schlapp T, Mercer AA;
PI Fleming SB;
XX
DR WPI; 2003-221750/21.
XX
PT New polynucleotide and recombinant proteins of Parapoxvirus ovis, useful
PT for manufacturing a medicament for treating virus related disease, viral
PT infections, non-viral infections, proliferative disease or inflammatory
PT disease.
XX
PS Example 4; Page 33; 51pp; English.
XX
CC The invention relates to a novel purified and isolated polynucleotide
CC (N1) of Parapoxvirus ovis (PPVO) comprising a nucleotide sequence (S1,
CC not defined in the specification), or its complementary sequence,
CC fragment or functional variant. A polynucleotide of the invention has
CC virucide, anti-HIV, hepatotropic, antiinflammatory, cytostatic,
CC vulnerary, antiasthmatic, antiallergic, dermatological, antidiabetic,
CC immunosuppressive, antirheumatic, antiarthritic, thyromimetic,
CC protozoacide, amoebicide, and antibacterial activity. The polynucleotides
CC may have a use in gene therapy. The recombinant proteins encoded by the
CC polynucleotides, or recombinant viruses comprising a Vaccinia virus
CC genome and fragments of a PPVO genome are useful for manufacturing
CC pharmaceutical compositions for treating virus related disease (e.g.
CC hepatitis, papillomatosis, herpes virus infections, liver fibrosis, HIV
CC infections or influenza), viral infections, non-viral infections (e.g.
CC infections with mycobacteria, mycoplasma, amoeba or plasmodia),
CC proliferative disease (e.g. cancer, leukaemia, warts or other skin
CC neoplasms), inflammatory disease (e.g. Crohn's disease, COPD, asthma or
CC conditions related to healing of wounds), allergic disease, and/or
CC autoimmune diseases (systemic lupus erythematosus, Sjogren's disease,
CC Hashimoto's thyroiditis, rheumatoid arthritis or diabetes mellitus). The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 7 AA;

Query Match 8.6%; Score 5; DB 7; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LQLLK 17
|||||
Db 2 LQLLK 6

RESULT 12
ADN08587
ID ADN08587 standard; peptide; 7 AA.
XX
AC ADN08587;
XX
DT 15-JUL-2004 (first entry)
XX
DE Ceramic tea stain binding peptide, SEQ ID No 152.
XX
KW tannin; polyphenolic; anthocyanin; tea; wine; stain; fabric; surface;
KW enzymatic; ceramic; hair; skin; detergent; cleaning; wine ageing;
KW fermentation; reducing; eliminating; astringency; conjugate; bleaching;
KW personal care product; cream; lotion; ointment.
XX
OS Unidentified.
XX
PN WO2004033482-A2.
XX
PD 22-APR-2004.
XX
PF 06-OCT-2003; 2003WO-US031776.
XX
PR 08-OCT-2002; 2002US-0417210P.
XX
PA (GEMV) GENENCOR INT INC.
XX
PI Murray CJ, Tijerina P, Van Gastel FJC;
XX

```
DR   WPI; 2004-340885/31.
XX
PT   Novel binding peptide binding to tannin, anthocyanin and phenolic
PT   compounds, useful for delivering agent to target, and for modifying tea
PT   or wine stain on fabric or surface chosen from ceramic, glasses, wood,
PT   paper, skin, hair and plastic.
XX
PS   Claim 1; SEQ ID NO 152; 38pp; English.
XX
CC   The invention relates to a novel peptide which binds to tannin,
CC   polyphenolic or anthocyanin compounds. The binding peptide consists
CC   essentially of any one of 316 amino acid sequences of 5-13 amino acids in
CC   length or has at least 70% identity to one of the 316 amino acid
CC   sequences. The binding peptide is useful for delivering an agent to a
CC   target, which involves conjugating the binding peptide to an agent to
CC   form a binding peptide conjugate and exposing a target to the binding
CC   peptide conjugate, where the binding peptide conjugate binds to the
CC   target. An enzymatic composition comprising the binding peptide is useful
CC   for modifying a tea or wine stain on a fabric or a surface, which
CC   involves contacting the surface having tea or wine stain with the
CC   enzymatic composition. The surface is preferably a ceramic surface, hair
CC   or skin. The binding peptide is useful for producing enzymatic
CC   compositions for use in detergent or cleaning compositions such as for
CC   removing food stains on fabrics or removing food stains on surfaces such
CC   as ceramic and teeth. The binding peptide is useful in the wine ageing
CC   process, where a tannin compound is targeted at the early stage of wine
CC   fermentation process for reducing or eliminating astringency of tannins
CC   in the wine. The binding conjugate, combined with a bleaching agent is
CC   useful for delivering a bleaching agent to stained teeth for bleaching.
CC   The binding peptide and binding conjugate are useful in personal care
CC   products such as creams, lotions and ointments. The binding peptide is
CC   useful for modifying astringency in skin. This sequence represents a
CC   binding peptide which specifically binds to tea stains on ceramics for
CC   use in the stain removal process of the invention.
XX
SQ   Sequence 7 AA;

  Query Match             8.6%;  Score 5;  DB 8;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 PPSPL 9
              |||||
Db          1 PPSPL 5


RESULT 13
ADP74892
ID   ADP74892 standard; peptide; 7 AA.
XX
AC   ADP74892;
XX
DT   23-SEP-2004  (first entry)
XX
DE   Parapoxvirus ovis genome ORF-encoding N-terminal peptide #42.
XX
KW   parapoxvirus ovis; PPVO; virucide; cytostatic; antiinflammatory;
KW   immunosuppressive; antiallergic; gene therapy; viral infection;
KW   non-viral infection; proliferative disease; inflammatory disease;
KW   allergic disease; autoimmune disease; open reading frame; ORF.
XX
OS   Parapoxvirus.
XX
PN   WO2004054614-A1.
XX
PD   01-JUL-2004.
XX
PF   17-DEC-2002; 2002WO-EP014402.
XX
PR   17-DEC-2002; 2002WO-EP014402.
XX
PA   (FARB ) BAYER HEALTHCARE AG.
XX
PI   Weber O,  Friederichs SM,  Siegling A,  Schlapp T,  Mercer AA;
PI   Fleming SB;
XX
DR   WPI; 2004-488011/46.
XX
PT   New polynucleotide coding for a parapoxvirus ovis (PPVO) genome, useful
PT   for preparing a composition for treating viral or non-viral infections,
PT   or proliferative, inflammatory, allergic or autoimmune diseases.
XX
PS   Example 4; Page 33; 103pp; English.
XX
CC   This invention relates to a novel purified and isolated polynucleotide
CC   which comprises a sequence having 137560 bp derived from the Parapoxvirus
CC   ovis (PPVO) and the open reading frames (ORFs) contained within it. The
CC   invention may be useful for the production of compounds with a virucide,
CC   cytostatic, antiinflammatory, immunosuppressive or antiallergic activity
CC   whilst the disclosed sequences may be useful for gene therapy. The
CC   invention is useful for preparing a composition for treating viral or non
CC   -viral infections, or proliferative, inflammatory, allergic or autoimmune
CC   diseases. The present sequence is that of a peptide which represents the
CC   N-terminal of a protein encoded by an ORF of the Parapoxvirus genome
CC   sequence of the invention.
XX
SQ   Sequence 7 AA;

  Query Match             8.6%;  Score 5;  DB 8;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          13 LQLLK 17
               |||||
Db           2 LQLLK 6


RESULT 14
AAB06282
ID   AAB06282 standard; peptide; 8 AA.
XX
AC   AAB06282;
XX
DT   03-OCT-2000  (first entry)
XX
DE   proSKI-1 putative zymogen cleavage site.
XX
KW   Human; pro-brain-derived neurotrophic factor; proBDNF;
KW   subtilisin-kexin isoenzyme 1; SKI-1; proSKI-1 zymogen processing;
KW   antilipaemic; cytostatic; vasotropic; SKI-1 inhibitor;
KW   hypercholesterolaemia; liver steatosis; Ras-dependent cancer; restenosis;
KW   amyloid protein formation.
XX
OS   Homo sapiens.
OS   Mus sp.
OS   Rattus sp.
XX
FH   Key             Location/Qualifiers
FT   Cleavage-site   4. .5
XX
PN   WO200026348-A2.
XX
PD   11-MAY-2000.
XX
PF   04-NOV-1999;   99WO-CA001058.
XX
PR   04-NOV-1998;   98CA-02249648.
XX
PA   (RECL-) INST RECH CLINIQUES MONTREAL.
XX
PI   Seidah N,  Chretien M,  Marcinkiewicz M,  Laaksonen R,  Davignon J;
XX
DR   WPI; 2000-365601/31.
XX
PT   Novel soluble proteic fragment of subtilisin-kexin isoenzyme for
PT   producing a polypeptide useful for treating hypercholesterolemia, liver
```

PT steatosis and amyloidosis, comprises a specific amino acid sequence.
XX
PS Example 1; Page 25; 119pp; English.
XX
CC The present sequence is a possible cleavage site in proSKI-1 which may be
CC targeted by a zymogen to form mature subtilisin-kexin isoenzyme 1 (SKI-
CC 1). The conserved site was identified by alignment of SKI-1 with other
CC subtilases. SKI-1 is a type-1 membrane-bound proteinase which cleaves pro
CC -brain-derived neurotrophic factor. Peptides which bind to and are
CC cleaved by SKI-1 may be used for monitoring SKI-1 activity, for screening
CC inhibitors of SKI-1 activity, or for screening enhancers of SKI-1
CC activity. Proteic fragments of SKI-1 which bind to the SKI-1 catalytic
CC site may be used as inhibitors of SKI-1 activity. They may be used to
CC treat diseases involving overexpression of SKI-1 or SKI-1 substrate. Such
CC diseases include hypercholesterolaemia, high levels of fatty acids,
CC lipids or farnesyl pyrophosphate, liver steatosis, Ras-dependent cancer,
CC restenosis and amyloid protein formation
XX
SQ Sequence 8 AA;

Query Match 8.6%; Score 5; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 SSRRL 23
|||||
Db 3 SSRRL 7

RESULT 15
AAR59128
ID AAR59128 standard; peptide; 9 AA.
XX
AC AAR59128;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAY-1995 (first entry)
XX
DE Peptide fragment (1.0882) of HBV binds HLA-A2.1.
XX
KW antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
KW core antigen; surface antigen; pharmaceutical composition; in vivo;
KW ex vivo; therapeutic; diagnostic; MHC class I molecule;
KW major histocompatability complex; HLA-A2.1; 9mer; 10mer; anchor;
KW human leukocyte antigen.
XX
OS Hepatitis B virus; strain adw.
XX
PN WO9420127-A1.
XX
PD 15-SEP-1994.
XX
PF 04-MAR-1994; 94WO-US002353.
XX
PR 05-MAR-1993; 93US-00027146.
PR 04-JUN-1993; 93US-00073205.
PR 29-NOV-1993; 93US-00159184.
XX
PA (CYTE-) CYTEL CORP.
XX
PI Grey HM, Sette A, Sidney J, Kast W;
XX
DR WPI; 1994-302678/37.
XX
PT Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
PT treatment or prophylaxis of cancer, virus infection or autoimmune
PT diseases.
XX
PS Example 5; Page 101; 138pp; English.
XX
CC AAR59104-264 are immunogenic 9mer peptides that contain a HLA-A2.1
CC binding motif. These peptides bind HLA-A2.1 and have a binding affinity
CC of at least 1% as compared to a reference peptide (AAR71293). AAR59128
CC has an IC50 of 0.046 and the sequence occurs at position 1088 in the HBV
CC POL protein. The peptides of the invention can induce cytotoxic T
CC lymphocytes which can react with target cells. They can be used for the
CC treatment or prophlyaxis of cancer, eg. prostate cancer or lymphoma, etc.
CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 9 AA;

Query Match 8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 SLMLL 38
|||||
Db 5 SLMLL 9

Search completed: October 16, 2005, 05:57:20
Job time : 119 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 16, 2005, 05:19:43 ; Search time 103.5 Seconds
(without alignments)
8269.939 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1813
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaatgaaaagtt 1027

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 3719576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135538_17068/app_query.fasta_1.1223
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10031158_@CGN_1_1_130_@runat_14102005_135538_17068
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1011	55.8	318	10	US-09-957-708-29	Sequence 29, Appl
2	1005	55.4	318	16	US-10-643-795A-103	Sequence 103, App
3	1005	55.4	318	17	US-10-948-518-103	Sequence 103, App
4	908	50.1	338	17	US-10-847-918-29	Sequence 29, Appl
5	798	44.0	274	14	US-10-340-536-30	Sequence 30, Appl
6	558	30.8	105	9	US-09-864-761-47901	Sequence 47901, A
7	483.5	26.7	101	9	US-09-925-300-998	Sequence 998, App
8	307	16.9	58	14	US-10-205-823-405	Sequence 405, App
9	307	16.9	58	18	US-10-756-149-5551	Sequence 5551, Ap
10	307	16.9	58	20	US-11-051-454-405	Sequence 405, App
c 11	196	10.8	534	18	US-10-450-763-35857	Sequence 35857, A
12	194	10.7	36	9	US-09-991-548-32	Sequence 32, Appl
13	138.5	7.6	244	18	US-10-450-763-49648	Sequence 49648, A
14	137.5	7.6	216	17	US-10-887-228A-6	Sequence 6, Appli
15	137.5	7.6	216	17	US-10-887-231-2	Sequence 2, Appli
16	137	7.6	686	15	US-10-072-012-795	Sequence 795, App
17	136.5	7.5	216	17	US-10-887-231-27	Sequence 27, Appl
18	134.5	7.4	214	16	US-10-128-520-107	Sequence 107, App
19	134.5	7.4	216	17	US-10-887-231-21	Sequence 21, Appl
20	133.5	7.4	215	16	US-10-128-520-102	Sequence 102, App
21	133.5	7.4	215	16	US-10-128-520-133	Sequence 133, App
22	133.5	7.4	216	16	US-10-128-520-114	Sequence 114, App
23	133.5	7.4	240	15	US-10-291-265-248	Sequence 248, App
24	133	7.3	218	16	US-10-379-392-159	Sequence 159, App
25	133	7.3	218	16	US-10-379-392-161	Sequence 161, App
26	133	7.3	218	16	US-10-379-392-164	Sequence 164, App
27	132.5	7.3	215	16	US-10-128-520-103	Sequence 103, App
28	132.5	7.3	216	17	US-10-887-231-23	Sequence 23, Appl
29	131.5	7.3	211	16	US-10-128-520-132	Sequence 132, App
30	131.5	7.3	231	15	US-10-291-265-283	Sequence 283, App
31	131	7.2	229	14	US-10-310-719-36	Sequence 36, Appl
32	130.5	7.2	211	16	US-10-128-520-99	Sequence 99, Appl
33	130.5	7.2	213	16	US-10-128-520-138	Sequence 138, App
34	130.5	7.2	214	16	US-10-128-520-122	Sequence 122, App
35	130.5	7.2	214	16	US-10-128-520-124	Sequence 124, App
36	130.5	7.2	214	16	US-10-128-520-127	Sequence 127, App
37	130.5	7.2	214	17	US-10-887-231-37	Sequence 37, Appl
38	130.5	7.2	215	16	US-10-128-520-135	Sequence 135, App
39	130.5	7.2	216	16	US-10-128-520-120	Sequence 120, App
40	130.5	7.2	216	17	US-10-887-228A-10	Sequence 10, Appl
41	130.5	7.2	216	17	US-10-887-231-10	Sequence 10, Appl
42	130.5	7.2	216	17	US-10-887-231-29	Sequence 29, Appl
43	130.5	7.2	216	17	US-10-887-231-31	Sequence 31, Appl
44	130.5	7.2	216	17	US-10-887-231-33	Sequence 33, Appl
45	130.5	7.2	216	17	US-10-887-231-39	Sequence 39, Appl

ALIGNMENTS

RESULT 1
US-09-957-708-29
; Sequence 29, Application US/09957708
; Publication No. US20030031678A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Yongming
; APPLICANT: Recipon, Herve
; APPLICANT: Cafferkey, Robert
; APPLICANT: Ali, Shujath
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific
; TITLE OF INVENTION: Genes
; FILE REFERENCE: DEX-0239
; CURRENT APPLICATION NUMBER: US/09/957,708
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/233,746
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 318

```
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-957-708-29

Alignment Scores:
Pred. No.:                7.24e-99        Length:         318
Score:                    1011.00         Matches:        191
Percent Similarity:       100.00%         Conservative:   1
Best Local Similarity:    99.48%          Mismatches:     0
Query Match:              55.76%          Indels:         0
DB:                       10              Gaps:           0

US-10-031-158-13 (1-1027) x US-09-957-708-29 (1-318)

Qy          4 CAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGAT 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GlnGluLeuGlyLysLysIleLysValPheGlyProGlyThrLysLeuIleIleThrAsp 146

Qy         64 AAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAA 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 LysGlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAlaGlu 166

Qy        124 ACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTT 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 ThrLysLeuGlnLysAlaGlyThrTyrLeuCysLeuLeuGluLysPhePheProAspVal 186

Qy        184 ATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAAC 243
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        187 IleLysIleHisTrpGlnGluLysLysSerAsnThrIleLeuGlySerGlnGluGlyAsn 206

Qy        244 ACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ThrMetLysThrAsnAspThrTyrMetLysPheSerTrpLeuThrValProGluLysSer 226

Qy        304 CTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAA 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuAspLysGluHisArgCysIleValArgHisGluAsnAsnLysAsnGlyValAspGln 246

Qy        364 GAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAATTGT 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GluIleIlePheProProIleLysThrAspValIleThrMetAspProLysAspAsnCys 266

Qy        424 TCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTACATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 SerLysAspAlaAsnAspThrLeuLeuLeuGlnLeuThrAsnThrSerAlaTyrTyrMet 286

Qy        484 TACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTGCTT 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 TyrLeuLeuLeuLeuLysSerValValTyrPheAlaIleIleThrCysCysLeuLeu 306

Qy        544 AGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCA 579
              ||||||||||||||||||||||||||||||||||||
Db        307 ArgArgThrAlaPheCysCysAsnGlyGluLysSer 318

RESULT 2
US-10-643-795A-103
; Sequence 103, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:
;  APPLICANT: FREDERIC J. DESAUVAGE
;  APPLICANT:  GRETCHEN FRANTZ
;  APPLICANT:  KENNETH J. HILLAN
;  APPLICANT:  PAUL POLAKIS
;  APPLICANT:  ANDREW POLSON
;  APPLICANT:  VICTORIA SMITH
;  APPLICANT:  SUSAN D. SPENCER
;  APPLICANT:  THOMAS D. WU
;  APPLICANT:  ZEMIN ZHANG
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF TUMOR
;  FILE REFERENCE: P5026R1-US
;  CURRENT APPLICATION NUMBER: US/10/643,795A
;  CURRENT FILING DATE:  2003-08-19
;  PRIOR APPLICATION NUMBER: US 60/404,809
;  PRIOR FILING DATE: 2002-08-19
;  PRIOR APPLICATION NUMBER: US 60/405,645
;  PRIOR FILING DATE: 2002-08-21
;  PRIOR APPLICATION NUMBER: US 60/413,192
;  PRIOR FILING DATE: 2002-09-23
;  PRIOR APPLICATION NUMBER: US 60/419,008
;  PRIOR FILING DATE: 2002-10-15
;  PRIOR APPLICATION NUMBER: US 60/426,847
;  PRIOR FILING DATE: 2002-11-15
;  PRIOR APPLICATION NUMBER: US 60/484,959
;  PRIOR FILING DATE: 2003-07-02
;  NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 103
;   LENGTH: 318
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-643-795A-103

Alignment Scores:
Pred. No.:                3.18e-98        Length:         318
Score:                    1005.00         Matches:        190
Percent Similarity:       99.48%          Conservative:   1
Best Local Similarity:    98.96%          Mismatches:     1
Query Match:              55.43%          Indels:         0
DB:                       16              Gaps:           0

US-10-031-158-13 (1-1027) x US-10-643-795A-103 (1-318)

Qy          4 CAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGAT 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GlnGluLeuGlyLysLysIleLysValPheGlyProGlyThrLysLeuIleIleThrAsp 146

Qy         64 AAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAA 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 LysGlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAlaGlu 166

Qy        124 ACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTT 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 ThrLysLeuGlnLysAlaGlyThrTyrLeuCysLeuLeuGluLysPhePheProAspVal 186

Qy        184 ATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAAC 243
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        187 IleLysIleHisTrpGlnGluLysLysSerAsnThrIleLeuGlySerGlnGluGlyAsn 206

Qy        244 ACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ThrMetLysThrAsnAspThrTyrMetLysPheSerTrpLeuThrValProGluLysSer 226

Qy        304 CTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAA 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuAspLysGluHisArgCysIleValArgHisGluAsnAsnLysAsnGlyValAspGln 246

Qy        364 GAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAATTGT 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GluIleIlePheProProIleLysThrAspValIleThrMetAspProLysAspAsnCys 266

Qy        424 TCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTACATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 SerLysAspAlaAsnAspThrLeuLeuLeuGlnLeuThrAsnThrSerAlaTyrTyrThr 286

Qy        484 TACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTGCTT 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 TyrLeuLeuLeuLeuLysSerValValTyrPheAlaIleIleThrCysCysLeuLeu 306

Qy        544 AGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCA 579
              ||||||||||||||||||||||||||||||||||||
Db        307 ArgArgThrAlaPheCysCysAsnGlyGluLysSer 318

RESULT 3
US-10-948-518-103
; Sequence 103, Application US/10948518
; Publication No. US20050064492A1
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 16, 2005, 04:57:57 ; Search time 37.5 Seconds
(without alignments)
4088.778 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1813
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaatgaaaagtt 1027

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135536_16979/app_query.fasta_1.1223
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=n2p.rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158_@CGN_1_1_33_@runat_14102005_135536_16979 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	917	50.6	323	6	5260223-4	Patent No. 5260223
2	917	50.6	323	6	5260223-4	Patent No. 5260223
3	798	44.0	274	1	US-08-256-964A-19	Sequence 19, Appl
4	407	22.4	202	6	5189147-6	Patent No. 5189147
5	407	22.4	202	6	5189147-6	Patent No. 5189147
6	287	15.8	53	4	US-09-513-999C-4756	Sequence 4756, Ap
7	194	10.7	36	3	US-09-028-937-32	Sequence 32, Appl
8	137.5	7.6	235	3	US-09-049-672A-10	Sequence 10, Appl
9	129.5	7.1	235	2	US-08-378-939-12	Sequence 12, Appl
10	127.5	7.0	238	3	US-08-793-450-6	Sequence 6, Appli
11	127.5	7.0	240	3	US-09-049-672A-11	Sequence 11, Appl
12	126.5	7.0	236	3	US-08-487-550-10	Sequence 10, Appl
13	126.5	7.0	236	4	US-09-526-098-10	Sequence 10, Appl
14	126.5	7.0	236	4	US-09-383-916-10	Sequence 10, Appl
15	125.5	6.9	233	3	US-08-523-894-6	Sequence 6, Appli
16	125.5	6.9	235	3	US-09-049-672A-12	Sequence 12, Appl
17	125.5	6.9	235	4	US-09-152-060-70	Sequence 70, Appl
18	125.5	6.9	235	4	US-09-152-060-88	Sequence 88, Appl
19	125.5	6.9	236	3	US-09-049-672A-7	Sequence 7, Appli
20	123.5	6.8	302	4	US-09-789-697A-21	Sequence 21, Appl
21	123.5	6.8	303	4	US-09-509-347-7	Sequence 7, Appli
22	122.5	6.8	234	3	US-08-487-550-2	Sequence 2, Appli
23	122.5	6.8	234	4	US-09-526-098-2	Sequence 2, Appli
24	122.5	6.8	234	4	US-09-383-916-2	Sequence 2, Appli
25	121.5	6.7	293	6	5189147-3	Patent No. 5189147
26	121.5	6.7	293	6	5189147-3	Patent No. 5189147
27	119.5	6.6	215	4	US-09-949-016-10696	Sequence 10696, A
28	118.5	6.5	145	4	US-09-949-016-8908	Sequence 8908, Ap
29	115.5	6.4	306	3	US-09-082-593-2	Sequence 2, Appli
30	114.5	6.3	234	4	US-09-372-425A-4	Sequence 4, Appli
31	114	6.3	236	4	US-09-315-926A-79	Sequence 79, Appl
32	113.5	6.3	200	6	5189147-10	Patent No. 5189147
33	113.5	6.3	200	6	5189147-10	Patent No. 5189147
34	113.5	6.3	234	4	US-09-472-087-17	Sequence 17, Appl
35	113.5	6.3	234	4	US-09-472-087-69	Sequence 69, Appl
36	113.5	6.3	237	1	US-08-398-612A-28	Sequence 28, Appl
37	113.5	6.3	237	1	US-08-398-611A-28	Sequence 28, Appl
38	113.5	6.3	237	2	US-08-491-334A-28	Sequence 28, Appl
39	113.5	6.3	237	3	US-09-027-449-25	Sequence 25, Appl
40	113.5	6.3	237	3	US-08-804-444A-25	Sequence 25, Appl
41	113.5	6.3	237	3	US-09-026-985-25	Sequence 25, Appl
42	113.5	6.3	237	4	US-09-121-952A-25	Sequence 25, Appl
43	113.5	6.3	237	4	US-09-234-340A-25	Sequence 25, Appl
44	113.5	6.3	238	1	US-08-398-613A-28	Sequence 28, Appl
45	113	6.2	224	4	US-09-456-090A-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
5260223-4
;Patent No. 5260223
; APPLICANT: BRENNER, MICHAEL B.;STROMINGER, JACK L.;SEIDMAN,
;JOHN G.;IP, STEPHEN H.;. KRANGEL, MICHAEL S.
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA
;T CELL RECEPTOR
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/187,698
; FILING DATE: 29-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,256
; FILING DATE: 29-OCT-1987
; APPLICATION NUMBER: 16,252
; FILING DATE: 19-FEB-1987
; APPLICATION NUMBER: 882,100
; FILING DATE: 03-JUL-1986
;SEQ ID NO:4:
; LENGTH: 323
5260223-4

Alignment Scores:

Pred. No.:	2.23e-98	Length:	323
Score:	917.00	Matches:	174
Percent Similarity:	90.05%	Conservative:	7
Best Local Similarity:	86.57%	Mismatches:	4
Query Match:	50.58%	Indels:	16
DB:	6	Gaps:	1

US-10-031-158-13 (1-1027) x 5260223-4 (1-323)

```
Qy         25 AAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAACAACTTGATGCAGATGTT 84
                 |||:::||||||   ||||||   |||::::::|||||||||||||||||||||||||
Db        123 LysLeuPheGlySerGlyThrThrLeuValValThrAspLysGlnLeuAspAlaAspVal 142
```

```
Qy         85 TCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAGAAGGCTGGA 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 SerProLysProThrIlePheLeuProSerIleAlaGluThrLysLeuGlnLysAlaGly 162

Qy        145 ACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATACATTGGGAAGAA 204
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||:::|||
Db        163 ThrTyrLeuCysLeuLeuGluLysPhePheProAspIleIleLysIleHisTrpGlnGlu 182

Qy        205 AAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACTAACGACACA 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 LysLysSerAsnThrIleLeuGlySerGlnGluGlyAsnThrMetLysThrAsnAspThr 202

Qy        265 TACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTGGACAAAGAACACAGATGT 324
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        203 TyrMetLysPheSerTrpLeuThrValProGluGluSerLeuAspLysGluHisArgCys 222

Qy        325 ATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAAATTATCTTTCCTCCAATA 384
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        223 IleValArgHisGluAsnAsnLysAsnGlyIleAspGlnGluIleIleProProIle 242

Qy        385 AAGACG----------------------------------------------------GATGTC 396
              ||||||                                                    ||||||
Db        243 LysThrAspValThrThrValAspProLysTyrAsnTyrSerLysAspAlaAsnAspVal 262

Qy        397 ATCACAATGGATCCCAAAGACAATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAG 456
              |||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        263 IleThrMetAspProLysAspAsnTrpSerLysAspAlaAsnAspThrLeuLeuLeuGln 282

Qy        457 CTCACAAACACCTCTGCATATTACATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTAT 516
              |||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        283 LeuThrAsnThrSerAlaTyrTyrThrTyrLeuLeuLeuLeuLysSerValValTyr 302

Qy        517 TTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAA 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 PheAlaIleIleThrCysCysLeuLeuArgArgThrAlaPheCysCysAsnGlyGluLys 322

Qy        577 TCA 579
              |||
Db        323 Ser 323
```

RESULT 2
5260223-4
;Patent No. 5260223
; APPLICANT: BRENNER, MICHAEL B.;STROMINGER, JACK L.;SEIDMAN,
;JOHN G.;IP, STEPHEN H.;. KRANGEL, MICHAEL S.
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA
;T CELL RECEPTOR
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/187,698
; FILING DATE: 29-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,256
; FILING DATE: 29-OCT-1987
; APPLICATION NUMBER: 16,252
; FILING DATE: 19-FEB-1987
; APPLICATION NUMBER: 882,100
; FILING DATE: 03-JUL-1986
;SEQ ID NO:4:
; LENGTH: 323
5260223-4

Alignment Scores:

Pred. No.:	2.23e-98	Length:	323
Score:	917.00	Matches:	174
Percent Similarity:	90.05%	Conservative:	7
Best Local Similarity:	86.57%	Mismatches:	4
Query Match:	50.58%	Indels:	16
DB:	6	Gaps:	1

US-10-031-158-13 (1-1027) x 5260223-4 (1-323)

```
Qy         25 AAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAACAACTTGATGCAGATGTT 84
              |||:::||||||   ||||||   |||:::::|||||||||||||||||||||||||||||
Db        123 LysLeuPheGlySerGlyThrThrLeuValValThrAspLysGlnLeuAspAlaAspVal 142

Qy         85 TCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAGAAGGCTGGA 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 SerProLysProThrIlePheLeuProSerIleAlaGluThrLysLeuGlnLysAlaGly 162

Qy        145 ACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATACATTGGGAAGAA 204
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||:::|||
Db        163 ThrTyrLeuCysLeuLeuGluLysPhePheProAspIleIleLysIleHisTrpGlnGlu 182

Qy        205 AAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACTAACGACACA 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 LysLysSerAsnThrIleLeuGlySerGlnGluGlyAsnThrMetLysThrAsnAspThr 202

Qy        265 TACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTGGACAAAGAACACAGATGT 324
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        203 TyrMetLysPheSerTrpLeuThrValProGluGluSerLeuAspLysGluHisArgCys 222

Qy        325 ATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAAATTATCTTTCCTCCAATA 384
              ||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        223 IleValArgHisGluAsnAsnLysAsnGlyIleAspGlnGluIleIleProProIle 242

Qy        385 AAGACG----------------------------------------------------GATGTC 396
              ||||||                                                    ||||||
Db        243 LysThrAspValThrThrValAspProLysTyrAsnTyrSerLysAspAlaAsnAspVal 262

Qy        397 ATCACAATGGATCCCAAAGACAATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAG 456
              |||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        263 IleThrMetAspProLysAspAsnTrpSerLysAspAlaAsnAspThrLeuLeuLeuGln 282

Qy        457 CTCACAAACACCTCTGCATATTACATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTAT 516
              |||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        283 LeuThrAsnThrSerAlaTyrTyrThrTyrLeuLeuLeuLeuLysSerValValTyr 302

Qy        517 TTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAA 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 PheAlaIleIleThrCysCysLeuLeuArgArgThrAlaPheCysCysAsnGlyGluLys 322

Qy        577 TCA 579
              |||
Db        323 Ser 323
```

RESULT 3
US-08-256-964A-19
; Sequence 19, Application US/08256964A
; Patent No. 5723309
; GENERAL INFORMATION:
; APPLICANT: BONNEVILLE, Marc
; TITLE OF INVENTION: PRODUCTION OF SUBUNITS OF SOLUBLE T
; TITLE OF INVENTION: RECEPTORS BY Co-TRANSFECTION, USES OF THE PRODUCTS THUS
; TITLE OF INVENTION: OBTAINED
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Young & Thompson
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,964A
; FILING DATE: 14-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92 14203
; FILING DATE: 25-NOV-1992
; ATTORNEY/AGENT INFORMATION:

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 15, 2005, 18:17:13 ; Search time 123.5 Seconds
(without alignments)
8516.685 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1813
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaaatgaaaagtt 1027

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135536_16953/app_query.fasta_1.1223
-DB=UniProt_03 -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158_@CGN_1_1_152_@runat_14102005_135536_16953 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1003	55.3	315	2	Q6IN96	Q6in96 homo sapien
2	987	54.4	191	2	Q6P5Q5	Q6p5q5 homo sapien
3	958	52.8	298	2	Q6INA4	Q6ina4 homo sapien
4	917	50.6	323	2	Q6PI92	Q6pi92 homo sapien
5	874	48.2	189	1	TCC_HUMAN	P03986 homo sapien
6	654.5	36.1	172	1	TCC2_MOUSE	P03985 mus musculu
7	645	35.6	198	2	Q7TN06	Q7tn06 mus musculu
8	630	34.7	169	1	TCC3_MOUSE	P06334 mus musculu
9	618	34.1	167	1	TCC1_MOUSE	P01853 mus musculu
10	593.5	32.7	340	2	Q6PKA9	Q6pka9 mus musculu
11	562.5	31.0	190	1	TCC4_MOUSE	P06335 mus musculu
12	144.5	8.0	235	2	Q99M11	Q99m11 mus musculu
13	133.5	7.4	130	2	Q9D8W4	Q9d8w4 mus musculu
14	130.5	7.2	234	2	Q6GMW3	Q6gmw3 homo sapien
15	130.5	7.2	235	2	Q6PJG0	Q6pjg0 homo sapien
16	129.5	7.1	235	2	Q6IN99	Q6in99 homo sapien
17	129.5	7.1	236	2	Q8NEJ1	Q8nej1 homo sapien
18	129.5	7.1	240	2	Q8WUK3	Q8wuk3 homo sapien
19	128.5	7.1	233	2	Q8N5F4	Q8n5f4 homo sapien
20	128.5	7.1	233	2	Q8TBC9	Q8tbc9 homo sapien
21	128.5	7.1	234	2	Q8N355	Q8n355 homo sapien
22	128.5	7.1	234	2	Q6GMV8	Q6gmv8 homo sapien
23	128.5	7.1	234	2	Q7Z2U7	Q7z2u7 homo sapien
24	128.5	7.1	236	2	Q6GMX3	Q6gmx3 homo sapien
25	128.5	7.1	236	2	Q6IPQ0	Q6ipq0 homo sapien
26	128.5	7.1	236	2	Q6PIQ7	Q6piq7 homo sapien
27	127.5	7.0	208	2	Q6V501	Q6v501 oryctolagus
28	126.5	7.0	230	2	Q7Z2U3	Q7z2u3 homo sapien
29	126.5	7.0	303	2	Q7TND8	Q7tnd8 mus musculu
30	125.5	6.9	236	2	Q6GMX4	Q6gmx4 homo sapien
31	124.5	6.9	235	2	Q6PIK1	Q6pik1 homo sapien
32	123.5	6.8	233	2	Q6GMW4	Q6gmw4 homo sapien
33	123.5	6.8	233	2	Q6NS96	Q6ns96 homo sapien
34	123.5	6.8	234	2	Q6NS95	Q6ns95 homo sapien
35	123.5	6.8	235	2	Q6GMW6	Q6gmw6 homo sapien
36	123.5	6.8	236	2	Q96E61	Q96e61 homo sapien
37	123.5	6.8	236	2	Q6GMV7	Q6gmv7 homo sapien
38	123.5	6.8	236	2	Q6P5S3	Q6p5s3 homo sapien
39	120.5	6.6	104	1	LAC1_RAT	P20766 rattus norv
40	119.5	6.6	233	2	Q96I69	Q96i69 homo sapien
41	118.5	6.5	235	2	Q6P2J1	Q6p2j1 homo sapien
42	118.5	6.5	237	2	Q6DHW4	Q6dhw4 homo sapien
43	117.5	6.5	213	1	ILL1_HUMAN	P15814 homo sapien
44	117.5	6.5	235	2	Q6PJF2	Q6pjf2 homo sapien
45	117.5	6.5	294	2	Q8K1Z5	Q8k1z5 mus musculu

ALIGNMENTS

RESULT 1
Q6IN96
ID Q6IN96 PRELIMINARY; PRT; 315 AA.
AC Q6IN96;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE TRGV9 protein.
GN Name=TRGV9;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;

```
RA   Strausberg R.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC072396; AAH72396.1; -.
DR   HSSP; P06323; 1AC6.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   InterPro; IPR003596; Ig_v.
DR   SMART; SM00409; IG; 1.
DR   SMART; SM00407; IGc1; 1.
DR   SMART; SM00406; IGv; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
SQ   SEQUENCE   315 AA;  35299 MW;  7A17E7130F84DB46 CRC64;

Alignment Scores:
Pred. No.:              2.22e-89        Length:         315
Score:                  1003.00         Matches:        190
Percent Similarity:     99.48%          Conservative:   0
Best Local Similarity:  99.48%          Mismatches:     1
Query Match:            55.32%          Indels:         0
DB:                     2               Gaps:           0

US-10-031-158-13 (1-1027) x Q6IN96 (1-315)

Qy          7 GAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAA 66
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 GluLeuGlyLysLysIleLysValPheGlyProGlyThrLysLeuIleIleThrAspLys 144

Qy         67 CAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACA 126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 GlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAlaGluThr 164

Qy        127 AAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATT 186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 LysLeuGlnLysAlaGlyThrTyrLeuCysLeuGluLysPhePheProAspValIle 184

Qy        187 AAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACC 246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 LysIleHisTrpGluGluLysLysSerAsnThrIleLeuGlySerGlnGluGlyAsnThr 204

Qy        247 ATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTG 306
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        205 MetLysThrAsnAspThrTyrMetLysPheSerTrpLeuThrValProGluLysSerLeu 224

Qy        307 GACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAA 366
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 AspLysGluHisArgCysIleValArgHisGluAsnAsnLysAsnGlyValAspGlnGlu 244

Qy        367 ATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAATTGTTCA 426
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 IleIlePheProProIleLysThrAspValIleThrMetAspProLysAspAsnCysSer 264

Qy        427 AAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTACATGTAC 486
              ||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        265 LysAspAlaAsnAspThrLeuLeuLeuGlnLeuThrAsnThrSerAlaTyrTyrThrTyr 284

Qy        487 CTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTGCTTAGA 546
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 LeuLeuLeuLeuLeuLysSerValValTyrPheAlaIleIleThrCysCysLeuLeuArg 304

Qy        547 AGAACGGCTTTCTGCTGCAATGGAGAGAAATCA 579
              ||||||||||||||||||||||||||||||||
Db        305 ArgThrAlaPheCysCysAsnGlyGluLysSer 315

RESULT 2
Q6P5Q5
ID   Q6P5Q5      PRELIMINARY;      PRT;   191 AA.
AC   Q6P5Q5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   TRGV9 protein (Fragment).
GN   Name=TRGV9;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Prostate;
RA   Strausberg R.;
RL   Submitted (NOV-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC062761; AAH62761.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003597; Ig_c1.
DR   SMART; SM00407; IGc1; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   191 AA;  21639 MW;  D503094A172F4F7B CRC64;

Alignment Scores:
Pred. No.:              7.49e-88        Length:         191
Score:                  987.00          Matches:        186
Percent Similarity:     100.00%         Conservative:   2
Best Local Similarity:  98.94%          Mismatches:     0
Query Match:            54.44%          Indels:         0
DB:                     2               Gaps:           0

US-10-031-158-13 (1-1027) x Q6P5Q5 (1-191)

Qy         16 AAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAACAACTTGAT 75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          4 LysLysIleLysValPheGlyProGlyThrLysLeuIleIleThrAspLysGlnLeuAsp 23

Qy         76 GCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAG 135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         24 AlaAspValSerProLysProThrIlePheLeuProSerIleAlaGluThrLysLeuGln 43

Qy        136 AAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATACAT 195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         44 LysAlaGlyThrTyrLeuCysLeuGluLysPhePheProAspValIleLysIleHis 63

Qy        196 TGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACT 255
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 TrpGlnGluLysLysSerAsnThrIleLeuGlySerGlnGluGlyAsnThrMetLysThr 83

Qy        256 AACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTGGACAAAGAA 315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 AsnAspThrTyrMetLysPheSerTrpLeuThrValProGluLysSerLeuAspLysGlu 103

Qy        316 CACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAAATTATCTTT 375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 16, 2005, 04:25:08 ; Search time 37 Seconds
(without alignments)
5341.331 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1813
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaaatgaaaagtt 1027

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135536_16965/app_query.fasta_1.1223
-DB=PIR_79 -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158_@CGN_1_1_38_@runat_14102005_135536_16965 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	917	50.6	323	2	S01895	T-cell receptor ga
2	916	50.5	173	2	A26659	T-cell receptor ga
3	910	50.2	340	2	S03517	T-cell receptor ga
4	874	48.2	189	1	RWHUGC	T-cell receptor ga
5	874	48.2	189	2	B26659	T-cell receptor ga
6	866	47.8	205	2	B26425	T-cell receptor ga
7	814	44.9	304	2	S04663	T-cell receptor ga
8	798	44.0	274	2	S33440	T-cell receptor ga
9	662.5	36.5	210	2	JC4683	T-cell receptor ga
10	654.5	36.1	172	1	RWMSC7	T-cell receptor ga
11	642.5	35.4	172	2	D26420	T-cell receptor ga
12	630	34.7	169	1	RWMSC2	T-cell receptor ga
13	628.5	34.7	171	2	S04831	T-cell receptor ga
14	618	34.1	167	1	RWMSC1	T-cell receptor ga
15	582.5	32.1	206	2	C49054	T-cell receptor ga
16	562.5	31.0	190	1	RWMS8C	T-cell receptor ga
17	562	31.0	204	2	S36287	T-cell receptor ga
18	553.5	30.5	234	2	B43546	T-cell receptor ga
19	551.5	30.4	246	2	A43546	T-cell receptor ga
20	550	30.3	179	2	S36289	T-cell receptor ga
21	549	30.3	216	2	S36288	T-cell receptor ga
22	270	14.9	52	2	A22340	T-cell receptor ga
23	270	14.9	202	2	S36293	T-cell receptor ga
24	267	14.7	168	2	S36294	T-cell receptor ga
25	263	14.5	210	2	S36297	T-cell receptor ga
26	259	14.3	203	2	S36291	T-cell receptor ga
27	253	14.0	203	2	S23043	T-cell receptor ga
28	241	13.3	185	2	S36290	T-cell receptor ga
29	204	11.3	179	2	S36295	T-cell receptor ga
30	145	8.0	218	2	A36198	T-cell receptor be
31	138.5	7.6	243	2	S25755	Ig lambda chain -
32	137.5	7.6	235	2	S25749	Ig lambda chain -
33	134.5	7.4	208	2	B49444	Ig lambda chain (N
34	133.5	7.4	214	2	PC4156	Ig lambda chain V
35	131.5	7.3	231	2	S25751	Ig lambda chain -
36	130.5	7.2	118	2	A39949	Ig lambda chain J-
37	130.5	7.2	226	2	S25745	Ig lambda chain -
38	129.5	7.1	216	2	S69130	Ig lambda chain (D
39	129.5	7.1	231	2	S25753	Ig lambda chain -
40	129.5	7.1	233	2	S25741	Ig lambda chain -
41	129.5	7.1	240	2	A41797	Ig light chain - s
42	128.5	7.1	118	2	S12442	Ig lambda chain (K
43	128.5	7.1	118	2	S12441	Ig lambda chain (K
44	128.5	7.1	190	2	S25740	Ig lambda chain -
45	128.5	7.1	213	2	S21066	Ig lambda chain V

ALIGNMENTS

RESULT 1
S01895
T-cell receptor gamma chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 21-Jan-2000
C;Accession: S01895
R;Hochstenbach, F.; Parker, C.; McLean, J.; Gieselmann, V.; Band, H.; Bank, I.; Chess, L
J. Exp. Med. 168, 761-776, 1988
A;Title: Characterization of a third form of the human T cell receptor gamma/delta.
A;Reference number: S01895; MUID:88316186; PMID:3261778
A;Accession: S01895
A;Molecule type: mRNA
A;Residues: 1-323 <HOC>
A;Cross-references: EMBL:Y00790; NID:g37017; PIDN:CAA68744.1; PID:g37018
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: glycoprotein; T-cell receptor
F;1-14/Domain: signal sequence #status predicted <SIG>
F;15-323/Product: T-cell receptor gamma chain #status predicted <MAT>
F;15-120/Region: V segment
F;121-134/Region: J segment
F;135-323/Domain: C region #status predicted <CRE>
F;159-224/Domain: immunoglobulin homology <IMM>
F;106,200,254,270,276,285/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 9.57e-80 Length: 323
Score: 917.00 Matches: 174
Percent Similarity: 90.05% Conservative: 7
Best Local Similarity: 86.57% Mismatches: 4
Query Match: 50.58% Indels: 16
DB: 2 Gaps: 1

US-10-031-158-13 (1-1027) x S01895 (1-323)

```
Qy         25 AAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAACAACTTGATGCAGATGTT 84
              |||:::||||||   ||||||   |||:::::|||||||||||||||||||||||||||
Db        123 LysLeuPheGlySerGlyThrThrLeuValValThrAspLysGlnLeuAspAlaAspVal 142
```

```
Qy         85 TCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAGAAGGCTGGA 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 SerProLysProThrIlePheLeuProSerIleAlaGluThrLysLeuGlnLysAlaGly 162

Qy        145 ACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATACATTGGGAAGAA 204
              ||||||||||||||||||||||||||||||||||:::|||||||||||||||||:::|||
Db        163 ThrTyrLeuCysLeuLeuGluLysPhePheProAspIleIleLysIleHisTrpGlnGlu 182

Qy        205 AAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACTAACGACACA 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 LysLysSerAsnThrIleLeuGlySerGlnGluGlyAsnThrMetLysThrAsnAspThr 202

Qy        265 TACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTGGACAAAGAACACAGATGT 324
              ||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        203 TyrMetLysPheSerTrpLeuThrValProGluGluSerLeuAspLysGluHisArgCys 222

Qy        325 ATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAAATTATCTTTCCTCCAATA 384
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        223 IleValArgHisGluAsnAsnLysAsnGlyIleAspGlnGluIleIleProProIle 242

Qy        385 AAGACG------------------------------------------------GATGTC 396
              ||||||                                                ||||||
Db        243 LysThrAspValThrThrValAspProLysTyrAsnTyrSerLysAspAlaAsnAspVal 262

Qy        397 ATCACAATGGATCCCAAAGACAATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAG 456
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        263 IleThrMetAspProLysAspAsnTrpSerLysAspAlaAsnAspThrLeuLeuLeuGln 282

Qy        457 CTCACAAACACCTCTGCATATTACATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTAT 516
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        283 LeuThrAsnThrSerAlaTyrTyrThrTyrLeuLeuLeuLeuLysSerValValTyr 302

Qy        517 TTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAA 576
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 PheAlaIleIleThrCysCysLeuLeuArgArgThrAlaPheCysCysAsnGlyGluLys 322

Qy        577 TCA 579
              |||
Db        323 Ser 323
```

RESULT 2
A26659
T-cell receptor gamma-1 chain C region - human
N;Alternate names: T-cell receptor gamma chain C region (D-PLL)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 21-Jan-2000
C;Accession: A26659; B22340
R;Lefranc, M.P.; Forster, A.; Rabbitts, T.H.
Proc. Natl. Acad. Sci. U.S.A. 83, 9596-9600, 1986
A;Title: Genetic polymorphism and exon changes of the constant regions of the human T-ce
A;Reference number: A94708; MUID:87092294; PMID:2879283
A;Accession: A26659
A;Molecule type: DNA
A;Residues: 1-173 <LEF>
R;Lefranc, M.P.; Rabbitts, T.H.
Nature 316, 464-466, 1985
A;Title: Two tandemly organized human genes encoding the T-cell gamma constant-region se
A;Reference number: A93365; MUID:85268055; PMID:2991773
A;Accession: B22340
A;Molecule type: DNA
A;Residues: 1-52 <LE2>
C;Genetics:
A;Introns: 111/1; 127/1
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: T-cell receptor
F;25-90/Domain: immunoglobulin homology <IMM>

Alignment Scores:

Pred. No.:	1.17e-79	Length:	173
Score:	916.00	Matches:	172
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.42%	Mismatches:	0
Query Match:	50.52%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-13 (1-1027) x A26659 (1-173)

```
Qy         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AspLysGlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAla 20

Qy        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GluThrLysLeuGlnLysAlaGlyThrTyrLeuCysLeuLeuGluLysPhePheProAsp 40

Qy        181 GTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240
              ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db         41 ValIleLysIleHisTrpGlnGluLysLysSerAsnThrIleLeuGlySerGlnGluGly 60

Qy        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AsnThrMetLysThrAsnAspThrTyrMetLysPheSerTrpLeuThrValProGluLys 80

Qy        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 SerLeuAspLysGluHisArgCysIleValArgHisGluAsnAsnLysAsnGlyValAsp 100

Qy        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 GlnGluIleIlePheProProIleLysThrAspValIleThrMetAspProLysAspAsn 120

Qy        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CysSerLysAspAlaAsnAspThrLeuLeuLeuGlnLeuThrAsnThrSerAlaTyrTyr 140

Qy        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 MetTyrLeuLeuLeuLeuLeuLysSerValValTyrPheAlaIleIleThrCysCysLeu 160

Qy        541 CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCA 579
              |||||||||||||||||||||||||||||||||||||||
Db        161 LeuArgArgThrAlaPheCysCysAsnGlyGluLysSer 173
```

RESULT 3
S03517
T-cell receptor gamma chain precursor (clone pM17c64) - human
C;Species: Homo sapiens (man)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jan-2000
C;Accession: S03517
R;Tighe, L.; Forster, A.; Clark, D.M.; Boylston, A.W.; Lavenir, I.; Rabbitts, T.H.
Eur. J. Immunol. 17, 1729-1736, 1987
A;Title: Unusual forms of T cell gamma mRNA in a human T cell leukemia cell line: implic
A;Reference number: S03517; MUID:88083067; PMID:2961573
A;Accession: S03517
A;Molecule type: mRNA
A;Residues: 1-340 <TIG>
A;Cross-references: EMBL:X06774; NID:g37346; PIDN:CAA29941.1; PID:g37347
A;Note: this sequence was determined from the differentiated gene
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-340/Product: T-cell receptor gamma chain #status predicted <MAT>
F;21-116/Domain: V region (V-gamma-8) #status predicted <VRE>
F;121-135/Domain: J region (J-gamma-2) #status predicted <JRE>
F;136-340/Domain: C region (C-gamma-2) #status predicted <CRE>
F;160-225/Domain: immunoglobulin homology <IMM>

Alignment Scores:

Pred. No.:	4.51e-79	Length:	340
Score:	910.00	Matches:	174
Percent Similarity:	83.41%	Conservative:	7
Best Local Similarity:	80.18%	Mismatches:	4
Query Match:	50.19%	Indels:	32
DB:	2	Gaps:	1

US-10-031-158-13 (1-1027) x S03517 (1-340)

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 15, 2005, 18:13:26 ; Search time 98 Seconds
(without alignments)
8106.182 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1813
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaatgaaaagtt 1027

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135535_16945/app_query.fasta_1.1223
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158_@CGN_1_1_154_@runat_14102005_135535_16945 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1011	55.8	318	5	AAE24064	Aae24064 Human pro
2	1005	55.4	318	8	ADL06523	Adl06523 Human tum
3	959	52.9	311	7	ADE61825	Ade61825 Human Pro
4	917	50.6	323	2	AAR32906	Aar32906 TCR gamma
5	916	50.5	221	8	ADR66006	Adr66006 Human pro
6	916	50.5	221	8	ADR66904	Adr66904 Human pro
7	911	50.2	345	7	ADE08976	Ade08976 Novel pro
8	908	50.1	338	8	ADP54940	Adp54940 Human PRO
9	907	50.0	324	7	ADE07898	Ade07898 Novel pro
10	894	49.3	324	1	AAP91885	Aap91885 Deduced s
11	798	44.0	274	2	AAR55705	Aar55705 Human lym
12	798	44.0	274	6	ABR82494	Abr82494 Human T c
13	798	44.0	275	7	ADE61829	Ade61829 Human Pro
14	658.5	36.3	331	7	ADE61823	Ade61823 Rat Prote
15	658.5	36.3	331	7	ADE61827	Ade61827 Rat Prote
16	649	35.8	304	1	AAP91816	Aap91816 Mammalian
17	649	35.8	304	2	AAW01533	Aaw01533 Cytotoxic
18	649	35.8	304	3	AAY50115	Aay50115 Murine cy
19	595	32.8	111	4	AAB66400	Aab66400 Human pro
20	558	30.8	105	4	ABB28899	Abb28899 Peptide #
21	558	30.8	105	4	AAM67238	Aam67238 Human bon
22	558	30.8	105	5	ABG36889	Abg36889 Human pep
23	483.5	26.7	101	3	AAB56420	Aab56420 Human pro
24	307	16.9	58	4	AAB66399	Aab66399 Human TCR
25	307	16.9	58	7	ADB75581	Adb75581 Prostate
26	307	16.9	58	8	ADP54863	Adp54863 Human PRO
27	293	16.2	67	8	ADP64625	Adp64625 Human onc
28	287	15.8	53	3	AAG00675	Aag00675 Human sec
29	242	13.3	57	1	AAP91813	Aap91813 Peptide d
30	242	13.3	57	3	AAY50116	Aay50116 Murine CT
c 31	196	10.8	534	4	ABG05498	Abg05498 Novel hum
32	194	10.7	36	2	AAY39432	Aay39432 Human T-c
33	194	10.7	36	5	AAU78032	Aau78032 Human T c
34	175	9.7	62	8	ADP64624	Adp64624 Human onc
35	138.5	7.6	244	4	ABG19289	Abg19289 Novel hum
36	137.5	7.6	235	4	AAB36212	Aab36212 Human imm
37	137	7.6	686	5	ADI17259	Adi17259 Human NOV
38	135.5	7.5	240	3	AAY96306	Aay96306 Human IGF
39	134.5	7.4	214	6	ABR01470	Abr01470 Human ant
40	133.5	7.4	214	2	AAW27091	Aaw27091 Mouse mon
41	133.5	7.4	215	6	ABR01496	Abr01496 Human ant
42	133.5	7.4	215	6	ABR01465	Abr01465 Human ant
43	133.5	7.4	216	6	ABR01477	Abr01477 Human ant
44	133.5	7.4	240	4	AAU14141	Aau14141 Human nov
45	132.5	7.3	215	6	ABR01466	Abr01466 Human ant

ALIGNMENTS

RESULT 1
AAE24064
ID AAE24064 standard; protein; 318 AA.
XX
AC AAE24064;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human prostate specific protein (PSP) #7.
XX
KW Human; prostate specific protein; PSP; prostate specific nucleic acid;
KW vaccine; transgenic; prostate cancer; gene therapy; transgenic animal;
KW PSNA.
XX
OS Homo sapiens.
XX
PN WO200224718-A1.
XX
PD 28-MAR-2002.
XX
PF 19-SEP-2001; 2001WO-US029386.
XX
PR 19-SEP-2000; 2000US-0233746P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Recipon H, Cafferkey R, Ali S;
XX
DR WPI; 2002-471216/50.
XX
PT Novel isolated prostate specific polypeptide useful for identifying,
PT diagnosing, monitoring, staging, imaging, and treating prostate cancer
PT and non-cancerous disease states in prostate.
XX

```
PS   Claim 37; Page 200-201; 210pp; English.
XX
CC   The invention relates to prostate specific proteins (PSP) and prostate
CC   specific nucleic acids (PSNA). Sequences of the invention are useful for
CC   identifying, diagnosing, monitoring, staging, imaging and treating
CC   prostate cancer and non-cancerous disease states in prostate. They are
CC   also useful for producing engineered prostate tissue for treatment and
CC   research. The PSNA sequences are used in gene therapy and for producing
CC   transgenic animals and cells. The invention is also used as vaccines. The
CC   present sequence is human prostate specific protein of the invention
XX
SQ   Sequence 318 AA;

Alignment Scores:
Pred. No.:                4.48e-102      Length:          318
Score:                    1011.00        Matches:         191
Percent Similarity:       100.00%        Conservative:    1
Best Local Similarity:    99.48%         Mismatches:      0
Query Match:              55.76%         Indels:          0
DB:                       5              Gaps:            0

US-10-031-158-13 (1-1027) x AAE24064 (1-318)

Qy          4 CAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGAT 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GlnGluLeuGlyLysLysIleLysValPheGlyProGlyThrLysLeuIleIleThrAsp 146

Qy         64 AAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAA 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 LysGlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAlaGlu 166

Qy        124 ACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTT 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 ThrLysLeuGlnLysAlaGlyThrTyrLeuCysLeuLeuGluLysPhePheProAspVal 186

Qy        184 ATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAAC 243
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        187 IleLysIleHisTrpGlnGluLysLysSerAsnThrIleLeuGlySerGlnGluGlyAsn 206

Qy        244 ACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ThrMetLysThrAsnAspThrTyrMetLysPheSerTrpLeuThrValProGluLysSer 226

Qy        304 CTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAA 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 LeuAspLysGluHisArgCysIleValArgHisGluAsnAsnLysAsnGlyValAspGln 246

Qy        364 GAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAATTGT 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 GluIleIleProProIleLysThrAspValIleThrMetAspProLysAspAsnCys 266

Qy        424 TCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTACATG 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 SerLysAspAlaAsnAspThrLeuLeuLeuGlnLeuThrAsnThrSerAlaTyrTyrMet 286

Qy        484 TACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTGCTT 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 TyrLeuLeuLeuLeuLeuLysSerValValTyrPheAlaIleIleThrCysCysLeuLeu 306

Qy        544 AGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCA 579
              ||||||||||||||||||||||||||||||||||||
Db        307 ArgArgThrAlaPheCysCysAsnGlyGluLysSer 318

RESULT 2
ADL06523
ID   ADL06523 standard; protein; 318 AA.
XX
AC   ADL06523;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human tumour-associated antigenic target (TAT) polypeptide #22.
XX
KW   Human; tumour-associated antigenic target; TAT; cell death; tumour;
KW   cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004016225-A2.
XX
PD   26-FEB-2004.
XX
PF   19-AUG-2003; 2003WO-US025892.
XX
PR   19-AUG-2002; 2002US-0404809P.
PR   21-AUG-2002; 2002US-0405645P.
PR   23-SEP-2002; 2002US-0413192P.
PR   15-OCT-2002; 2002US-0419008P.
PR   15-NOV-2002; 2002US-0426847P.
PR   02-JUL-2003; 2003US-0484959P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Desauvage FJ,  Frantz G,  Hillan KJ,  Polakis P,  Polson A,  Smith V;
PI   Spencer SD,  Wu TD,  Zhang Z;
XX
DR   WPI; 2004-257144/24.
DR   N-PSDB; ADL06443.
XX
PT   New antibody that binds to a tumor-associated antigenic target (TAT)
PT   polypeptide, useful for preparing a composition for diagnosing or
PT   treating cancer.
XX
PS   Claim 2; SEQ ID NO 103; 319pp; English.
XX
CC   The present invention relates to the isolation of human tumour-associated
CC   antigenic target (TAT) polynucleotide and polypeptide sequences. Also
CC   disclosed is an antibody that binds to a TAT polypeptide. The antibody is
CC   a monoclonal antibody, an antibody fragment, a chimeric antibody or a
CC   humanised antibody. It is conjugated to a growth inhibitory agent. It is
CC   produced in bacteria or in CHO cells and induces death of a cell to which
CC   it binds. The antibody is useful for preparing a composition for
CC   diagnosing or treating tumours and cancer. The present sequence
CC   represents a human TAT polypeptide of the invention.
XX
SQ   Sequence 318 AA;

Alignment Scores:
Pred. No.:                2.06e-101      Length:          318
Score:                    1005.00        Matches:         190
Percent Similarity:       99.48%         Conservative:    1
Best Local Similarity:    98.96%         Mismatches:      1
Query Match:              55.43%         Indels:          0
DB:                       8              Gaps:            0

US-10-031-158-13 (1-1027) x ADL06523 (1-318)

Qy          4 CAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGAT 63
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GlnGluLeuGlyLysLysIleLysValPheGlyProGlyThrLysLeuIleIleThrAsp 146

Qy         64 AAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAA 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 LysGlnLeuAspAlaAspValSerProLysProThrIlePheLeuProSerIleAlaGlu 166

Qy        124 ACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTT 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        167 ThrLysLeuGlnLysAlaGlyThrTyrLeuCysLeuLeuGluLysPhePheProAspVal 186

Qy        184 ATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAAC 243
              ||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db        187 IleLysIleHisTrpGlnGluLysLysSerAsnThrIleLeuGlySerGlnGluGlyAsn 206

Qy        244 ACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCA 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ThrMetLysThrAsnAspThrTyrMetLysPheSerTrpLeuThrValProGluLysSer 226
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 15, 2005, 17:17:31 ; Search time 830 Seconds
(without alignments)
8673.699 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1027
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaatgaaaagtt 1027

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 8765912 seqs, 3504951483 residues

Total number of hits satisfying chosen parameters: 17531824

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1025.4	99.8	1027	15	US-10-205-823-404	Sequence 404, App
2	1025.4	99.8	1027	22	US-10-756-149-2859	Sequence 2859, Ap
3	1025.4	99.8	1027	24	US-11-051-454-404	Sequence 404, App
4	1012.2	98.6	1395	20	US-10-643-795A-23	Sequence 23, Appl
5	1012.2	98.6	1395	21	US-10-948-518-23	Sequence 23, Appl
6	966.6	94.1	1799	15	US-10-101-510-492	Sequence 492, App
7	944	91.9	1418	10	US-09-957-708-16	Sequence 16, Appl
8	906	88.2	1155	9	US-09-925-300-58	Sequence 58, Appl
9	900.6	87.7	1421	9	US-09-954-456-317	Sequence 317, App
10	900.6	87.7	1421	21	US-10-843-641A-3344	Sequence 3344, Ap
11	836.4	81.4	1586	10	US-09-960-706-676	Sequence 676, App
12	836.4	81.4	1586	10	US-09-873-319-424	Sequence 424, App
13	836.4	81.4	1586	21	US-10-847-918-14	Sequence 14, Appl
c 14	795.8	77.5	1162	20	US-10-357-930-21517	Sequence 21517, A
c 15	795.8	77.5	1162	20	US-10-357-930-24142	Sequence 24142, A
c 16	795.8	77.5	1162	20	US-10-357-930-24529	Sequence 24529, A
c 17	795.8	77.5	1162	20	US-10-357-930-24533	Sequence 24533, A
c 18	795.8	77.5	1162	20	US-10-357-930-24840	Sequence 24840, A
c 19	795.8	77.5	1162	20	US-10-357-930-25151	Sequence 25151, A
c 20	795.8	77.5	1162	20	US-10-357-930-27359	Sequence 27359, A
c 21	795.8	77.5	1162	20	US-10-357-930-27740	Sequence 27740, A
c 22	619	60.3	924	20	US-10-357-930-25081	Sequence 25081, A
23	589	57.4	124933	22	US-10-737-082-97	Sequence 97, Appl
24	589	57.4	124933	22	US-10-765-790-97	Sequence 97, Appl
c 25	497.2	48.4	1316	20	US-10-357-930-25082	Sequence 25082, A
c 26	497.2	48.4	1316	20	US-10-357-930-26474	Sequence 26474, A
c 27	471.6	45.9	486	20	US-10-357-930-25002	Sequence 25002, A
28	450.8	43.9	825	16	US-10-340-536-29	Sequence 29, Appl
29	446.8	43.5	784	20	US-10-357-930-21887	Sequence 21887, A
c 30	444	43.2	511	9	US-09-759-143-72	Sequence 72, Appl
c 31	444	43.2	511	9	US-09-780-669-72	Sequence 72, Appl
c 32	444	43.2	511	9	US-09-030-606-72	Sequence 72, Appl
c 33	444	43.2	511	9	US-09-822-827-72	Sequence 72, Appl
c 34	444	43.2	511	9	US-09-115-453-72	Sequence 72, Appl
c 35	444	43.2	511	9	US-09-232-880-72	Sequence 72, Appl
c 36	444	43.2	511	9	US-09-895-793-72	Sequence 72, Appl
c 37	444	43.2	511	9	US-09-895-814-72	Sequence 72, Appl
c 38	444	43.2	511	13	US-10-012-896-72	Sequence 72, Appl
c 39	444	43.2	511	14	US-10-010-940-72	Sequence 72, Appl
c 40	444	43.2	511	16	US-10-144-678A-72	Sequence 72, Appl
c 41	444	43.2	511	16	US-10-294-025-72	Sequence 72, Appl
c 42	444	43.2	511	19	US-10-688-838-72	Sequence 72, Appl
c 43	414.6	40.4	569	20	US-10-357-930-43494	Sequence 43494, A
c 44	408.8	39.8	539	20	US-10-357-930-22784	Sequence 22784, A
c 45	408.8	39.8	539	20	US-10-357-930-28636	Sequence 28636, A

ALIGNMENTS

RESULT 1
US-10-205-823-404
; Sequence 404, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12

```
;  PRIOR APPLICATION NUMBER: 60/362,158
;  PRIOR FILING DATE: 2002-03-05
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 404
;   LENGTH: 1027
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-205-823-404

  Query Match             99.8%;  Score 1025.4;  DB 15;  Length 1027;
  Best Local Similarity   99.9%;  Pred. No. 4.5e-294;
  Matches 1026;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60

Qy         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120

Qy        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180

Qy        181 GTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 GTTATTAAGATACATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240

Qy        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300

Qy        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360

Qy        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420

Qy        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480

Qy        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540

Qy        541 CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGA 600

Qy        601 GGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660

Qy        661 TTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTA 720

Qy        721 TAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT 780

Qy        781 AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCC 840

Qy        841 AAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTG 900

Qy        901 TTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960

Qy        961 TTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAATG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAATG 1020

Qy       1021 AAAAGTT 1027
              |||||||
Db       1021 AAAAGTT 1027


RESULT 2
US-10-756-149-2859
; Sequence 2859, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2859
;   LENGTH: 1027
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
US-10-756-149-2859

  Query Match             99.8%;  Score 1025.4;  DB 22;  Length 1027;
  Best Local Similarity   99.9%;  Pred. No. 4.5e-294;
  Matches 1026;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60

Qy         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120

Qy        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180

Qy        181 GTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        181 GTTATTAAGATACATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240

Qy        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300

Qy        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360

Qy        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420

Qy        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480

Qy        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 15, 2005, 15:48:55 ; Search time 221 Seconds
(without alignments)
7603.874 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1027
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaatgaaaagtt 1027

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	514.8	50.1	1080	6	5260223-3	Patent No. 5260223
	2	514.8	50.1	1080	6	5260223-3	Patent No. 5260223
	3	450.8	43.9	825	1	US-08-256-964A-18	Sequence 18, Appl
c	4	444	43.2	511	3	US-09-020-956-72	Sequence 72, Appl
c	5	444	43.2	511	3	US-09-030-607-72	Sequence 72, Appl
c	6	444	43.2	511	3	US-09-439-313-72	Sequence 72, Appl
c	7	444	43.2	511	3	US-09-352-616A-72	Sequence 72, Appl
c	8	444	43.2	511	3	US-09-232-149A-72	Sequence 72, Appl
c	9	444	43.2	511	4	US-09-159-812-72	Sequence 72, Appl
c	10	444	43.2	511	4	US-09-636-215-72	Sequence 72, Appl
c	11	444	43.2	511	4	US-09-685-166A-72	Sequence 72, Appl
c	12	444	43.2	511	4	US-09-115-453-72	Sequence 72, Appl
c	13	444	43.2	511	4	US-09-688-489-72	Sequence 72, Appl
c	14	444	43.2	511	4	US-09-679-426-72	Sequence 72, Appl
c	15	444	43.2	511	4	US-09-759-143-72	Sequence 72, Appl
c	16	444	43.2	511	4	US-09-651-236-72	Sequence 72, Appl
	17	407	39.6	467	3	US-09-020-956-75	Sequence 75, Appl
	18	407	39.6	467	3	US-09-030-607-75	Sequence 75, Appl
	19	407	39.6	467	3	US-09-439-313-75	Sequence 75, Appl
	20	407	39.6	467	3	US-09-352-616A-75	Sequence 75, Appl
	21	407	39.6	467	3	US-09-232-149A-75	Sequence 75, Appl
	22	407	39.6	467	4	US-09-159-812-75	Sequence 75, Appl
	23	407	39.6	467	4	US-09-636-215-75	Sequence 75, Appl
	24	407	39.6	467	4	US-09-685-166A-75	Sequence 75, Appl
	25	407	39.6	467	4	US-09-115-453-75	Sequence 75, Appl
	26	407	39.6	467	4	US-09-688-489-75	Sequence 75, Appl
	27	407	39.6	467	4	US-09-679-426-75	Sequence 75, Appl
	28	407	39.6	467	4	US-09-759-143-75	Sequence 75, Appl
	29	407	39.6	467	4	US-09-651-236-75	Sequence 75, Appl
	30	402.8	39.2	426	4	US-09-513-999C-679	Sequence 679, App
c	31	250.6	24.4	281	3	US-09-439-313-433	Sequence 433, App
c	32	250.6	24.4	281	3	US-09-352-616A-433	Sequence 433, App
c	33	250.6	24.4	281	4	US-09-636-215-433	Sequence 433, App
c	34	250.6	24.4	281	4	US-09-685-166A-433	Sequence 433, App
c	35	250.6	24.4	281	4	US-09-679-426-433	Sequence 433, App
c	36	250.6	24.4	281	4	US-09-759-143-433	Sequence 433, App
c	37	250.6	24.4	281	4	US-09-651-236-433	Sequence 433, App
	38	249.2	24.3	460	4	US-09-513-999C-8197	Sequence 8197, Ap
	39	95.2	9.3	223	4	US-09-636-215-673	Sequence 673, App
	40	95.2	9.3	223	4	US-09-685-166A-673	Sequence 673, App
	41	95.2	9.3	223	4	US-09-679-426-673	Sequence 673, App
	42	95.2	9.3	223	4	US-09-759-143-673	Sequence 673, App
	43	95.2	9.3	223	4	US-09-651-236-673	Sequence 673, App
c	44	72.8	7.1	214	4	US-09-636-215-671	Sequence 671, App
c	45	72.8	7.1	214	4	US-09-685-166A-671	Sequence 671, App

ALIGNMENTS

RESULT 1
5260223-3
;Patent No. 5260223
; APPLICANT: BRENNER, MICHAEL B.;STROMINGER, JACK L.;SEIDMAN,
;JOHN G.;IP, STEPHEN H.;. KRANGEL, MICHAEL S.
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA
;T CELL RECEPTOR
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/187,698
; FILING DATE: 29-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,256
; FILING DATE: 29-OCT-1987
; APPLICATION NUMBER: 16,252
; FILING DATE: 19-FEB-1987
; APPLICATION NUMBER: 882,100
; FILING DATE: 03-JUL-1986
;SEQ ID NO:3:
; LENGTH: 1080
5260223-3

Query Match 50.1%; Score 514.8; DB 6; Length 1080;
Best Local Similarity 88.7%; Pred. No. 9.9e-154;
Matches 589; Conservative 0; Mismatches 27; Indels 48; Gaps 1;

```
Qy          25 AAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAACAACTTGATGCAGATGTT 84
               ||  | |||||     |||||||  ||| |  | ||||||||||||||||||||||||||||
Db         403 AAACTCTTTGGCAGTGGAACAACACTTGTTGTCACAGATAAACAACTTGATGCAGATGTT 462

Qy          85 TCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAGAAGGCTGGA 144
               |||||||||||||||||||||||||||||| |||||||||||||| ||||||||||||||
Db         463 TCCCCCAAGCCCACTATTTTTCTTCCTTCGATTGCTGAAACAAAACTCCAGAAGGCTGGA 522

Qy         145 ACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATACATTGGGAAGAA 204
               |||||||||||||||||||||||||||||||| ||| |||||||||||| ||||| |||||
Db         523 ACATACCTTTGTCTTCTTGAGAAATTTTTCCCAGATATTATTAAGATAGATTGGCAAGAA 582

Qy         205 AAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACTAACGACACA 264
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         583 AAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACTAACGACACA 642

Qy         265 TACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTGGACAAAGAACACAGATGT 324
               |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db         643 TACATGAAATTTAGCTGGTTAACGGTGCCAGAAGAGTCACTGGACAAAGAACACAGATGT 702

Qy         325 ATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAAATTATCTTTCCTCCAATA 384
               |||||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||||
Db         703 ATCGTCAGACATGAGAATAATAAAAACGGAATTGATCAAGAAATTATCTTTGCTCCAATA 762
```

```
Qy        385 AAGAC----------------------------------------------------GGATGTC 396
              |||||                                                    |||||||
Db        763 AAGACAGATGTCACCACAGTGGATCCCAAATACAATTATTCAAAGGATGCAAATGATGTC 822

Qy        397 ATCACAATGGATCCCAAAGACAATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAG 456
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        823 ATCACAATGGATCCCAAAGACAATTGGTCAAAAGATGCAAATGATACACTACTGCTGCAG 882

Qy        457 CTCACAAACACCTCTGCATATTACATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTAT 516
              |||| ||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        883 CTCAGAAACACCTCTGCATATTACACGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTAT 942

Qy        517 TTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAA 576
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        943 TTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACTGCTTTCTGCTGCAATGGAGAGAAA 1002

Qy        577 TCATAACAGACGGTGGCACAAGGAGGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAG 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 TCATAACAGACGGTGGCACAAGGAGGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAG 1062

Qy        637 CGTC 640
              ||||
Db       1063 CGTC 1066
```

RESULT 2
5260223-3
;Patent No. 5260223
; APPLICANT: BRENNER, MICHAEL B.;STROMINGER, JACK L.;SEIDMAN,
;JOHN G.;IP, STEPHEN H.;. KRANGEL, MICHAEL S.
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA
;T CELL RECEPTOR
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/187,698
;   FILING DATE: 29-APR-1988
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 115,256
;   FILING DATE: 29-OCT-1987
;   APPLICATION NUMBER: 16,252
;   FILING DATE: 19-FEB-1987
;   APPLICATION NUMBER: 882,100
;   FILING DATE: 03-JUL-1986
;SEQ ID NO:3:
; LENGTH: 1080
5260223-3

  Query Match   50.1%; Score 514.8; DB 6; Length 1080;
  Best Local Similarity 88.7%; Pred. No. 9.9e-154;
  Matches 589; Conservative   0; Mismatches   27; Indels   48; Gaps   1;

```
Qy         25 AAGGTATTTGGTCCCGGAACAAAGCTTATCATTACAGATAAACAACTTGATGCAGATGTT 84
              ||  | |||||    |||||||  ||| |  | ||||||||||||||||||||||||||
Db        403 AAACTCTTTGGCAGTGGAACAACACTTGTTGTCACAGATAAACAACTTGATGCAGATGTT 462

Qy         85 TCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCCAGAAGGCTGGA 144
              |||||||||||||||||||||||||| |||||||||||| |||||||||||||||||
Db        463 TCCCCCAAGCCCACTATTTTTCTTCCTTCGATTGCTGAAACAAAACTCCAGAAGGCTGGA 522

Qy        145 ACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATACATTGGGAAGAA 204
              |||||||||||||||||||||||||||||| ||| |||||||||| ||||| |||||
Db        523 ACATACCTTTGTCTTCTTGAGAAATTTTTCCCAGATATTATTAAGATAGATTGGCAAGAA 582

Qy        205 AAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACTAACGACACA 264
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        583 AAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCATGAAGACTAACGACACA 642

Qy        265 TACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAGTCACTGGACAAAGAACACAGATGT 324
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db        643 TACATGAAATTTAGCTGGTTAACGGTGCCAGAAGAGTCACTGGACAAAGAACACAGATGT 702

Qy        325 ATCGTCAGACATGAGAATAATAAAAACGGAGTTGATCAAGAAATTATCTTTCCTCCAATA 384
              |||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||||
```

```
Db        703 ATCGTCAGACATGAGAATAATAAAAACGGAATTGATCAAGAAATTATCTTTGCTCCAATA 762

Qy        385 AAGAC----------------------------------------------------GGATGTC 396
              |||||                                                    |||||||
Db        763 AAGACAGATGTCACCACAGTGGATCCCAAATACAATTATTCAAAGGATGCAAATGATGTC 822

Qy        397 ATCACAATGGATCCCAAAGACAATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAG 456
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        823 ATCACAATGGATCCCAAAGACAATTGGTCAAAAGATGCAAATGATACACTACTGCTGCAG 882

Qy        457 CTCACAAACACCTCTGCATATTACATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTAT 516
              |||| ||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        883 CTCAGAAACACCTCTGCATATTACACGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTAT 942

Qy        517 TTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAA 576
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        943 TTTGCCATCATCACCTGCTGTCTGCTTAGAAGAACTGCTTTCTGCTGCAATGGAGAGAAA 1002

Qy        577 TCATAACAGACGGTGGCACAAGGAGGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAG 636
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1003 TCATAACAGACGGTGGCACAAGGAGGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAG 1062

Qy        637 CGTC 640
              ||||
Db       1063 CGTC 1066
```

RESULT 3
US-08-256-964A-18
; Sequence 18, Application US/08256964A
; Patent No. 5723309
; GENERAL INFORMATION:
;   APPLICANT: BONNEVILLE, Marc
;   TITLE OF INVENTION: PRODUCTION OF SUBUNITS OF SOLUBLE T
;   TITLE OF INVENTION: RECEPTORS BY Co-TRANSFECTION, USES OF THE PRODUCTS THUS
;   TITLE OF INVENTION: OBTAINED
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Young & Thompson
;   STREET: 745 South 23rd Street
;   CITY: Arlington
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/256,964A
;   FILING DATE: 14-SEP-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: FR 92 14203
;   FILING DATE: 25-NOV-1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: PATCH, Andrew J.
;   REGISTRATION NUMBER: Reg. No. 5723309 32,925
;   REFERENCE/DOCKET NUMBER: BE 94/449
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703/521-2297
;   TELEFAX: 703/685-0573
;   TELEX: 248425
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 825 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 15, 2005, 15:47:50 ; Search time 3341 Seconds
(without alignments)
11700.690 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1027
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaatgaaaagtt 1027

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	963.4	93.8	3533	3	BC030554	BC030554 Homo sapi
	2	879.2	85.6	1510	3	BC017861	BC017861 Homo sapi
	3	840.8	81.9	1757	3	BC039116	BC039116 Homo sapi
	4	836.2	81.4	914	5	BQ934484	BQ934484 AGENCOURT
	5	634.4	61.8	821	2	BF679123	BF679123 602153390
	6	627.8	61.1	997	2	BF678971	BF678971 602153608
c	7	618.6	60.2	636	1	AI972955	AI972955 wr45g05.x
c	8	613.8	59.8	638	1	AI685999	AI685999 tt91a04.x
	9	612.4	59.6	902	2	BF674593	BF674593 602137811
	10	603.6	58.8	959	7	CN645427	CN645427 ILLUMIGEN
	11	590.2	57.5	703	4	BG217853	BG217853 RST37574
c	12	589.6	57.4	629	1	AI768834	AI768834 wj03h04.x
c	13	584	56.9	700	6	CA449324	CA449324 UI-H-EI1-
c	14	580.4	56.5	582	2	BE326754	BE326754 hr64d05.x
	15	578.8	56.4	613	4	BI905966	BI905966 603062856
	16	571.4	55.6	719	2	BF681238	BF681238 602155502
	17	566.2	55.1	670	2	BF677916	BF677916 602084736
	18	564	54.9	789	5	BU199143	BU199143 DCBCIA08
	19	559.4	54.5	629	2	BF679165	BF679165 602153737
	20	556.8	54.2	1058	4	BM544213	BM544213 AGENCOURT
	21	554.6	54.0	881	2	BF677648	BF677648 602085529
	22	551	53.7	849	2	BF679284	BF679284 602153477
c	23	547.6	53.3	701	9	AG175832	AG175832 Pan trogl
	24	547.2	53.3	803	2	BF680840	BF680840 602156088
	25	542	52.8	857	2	BF674457	BF674457 602137231
	26	538.6	52.4	861	2	BF681385	BF681385 602156579
c	27	532.4	51.8	665	1	AA569813	AA569813 nm46c02.s
c	28	530.8	51.7	600	1	AI244217	AI244217 qj86c05.x
	29	517.6	50.4	1055	4	BF964904	BF964904 602268596
	30	491.8	47.9	927	2	BF679476	BF679476 602153280
c	31	490.4	47.8	503	1	AI253507	AI253507 aq33c11.x
	32	490	47.7	525	5	BX112931	BX112931 BX112931
	33	489.2	47.6	572	2	BF370026	BF370026 MR3-FN000
c	34	488.4	47.6	513	1	AI823897	AI823897 wj28e10.x
c	35	466	45.4	479	2	AW575113	AW575113 UI-HF-BK0
c	36	448.6	43.7	529	1	AA908306	AA908306 og33h07.s
	37	447.8	43.6	596	1	AV715641	AV715641 AV715641
c	38	435.6	42.4	442	1	AI718234	AI718234 as42h05.x
c	39	434.8	42.3	497	1	AA654656	AA654656 nt76c02.s
	40	430	41.9	870	2	BF678129	BF678129 602085181
c	41	429.2	41.8	512	1	AA657507	AA657507 nt64g07.s
c	42	428	41.7	460	2	BE772783	BE772783 RC1-FT013
	43	427	41.6	681	1	AV714756	AV714756 AV714756
	44	418.8	40.8	865	1	AI557112	AI557112 PT2.1_13_
	45	417	40.6	523	2	BF675605	BF675605 602083495

ALIGNMENTS

RESULT 1
BC030554
LOCUS BC030554 3533 bp mRNA linear HTC 20-MAY-2002
DEFINITION Homo sapiens, Similar to T cell receptor gamma locus, clone IMAGE:5213435, mRNA.
ACCESSION BC030554
VERSION BC030554.1 GI:20988582
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3533)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 64 Row: k Column: 18
This clone has the following problem: frame shifted.
FEATURES Location/Qualifiers
source 1. .3533
/organism="Homo sapiens"
/mol_type="mRNA"

```
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5213435"
                     /tissue_type="Blood, adult leukocytes"
                     /clone_lib="NIH_MGC_118"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
ORIGIN

  Query Match             93.8%;  Score 963.4;  DB 3;  Length 3533;
  Best Local Similarity   99.5%;  Pred. No. 2.5e-253;
  Matches  964;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy         59 CAGATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTG 118
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 CTGATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTG 802

Qy        119 CTGAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTG 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        803 CTGAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTG 862

Qy        179 ATGTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGG 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 ATGTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGG 922

Qy        239 GGAACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAA 298
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        923 GGAACACCATGAAGACTAATGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAA 982

Qy        299 AGTCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTG 358
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        983 AGTCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTG 1042

Qy        359 ATCAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACA 418
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1043 ATCAAGAAATTATCTTTCCTCCAATAAAGACAGATGTCATCACAATGGATCCCAAAGACA 1102

Qy        419 ATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATT 478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1103 ATTGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATT 1162

Qy        479 ACATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTC 538
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1163 ACACGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTC 1222

Qy        539 TGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAG 598
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1223 TGCTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAG 1282

Qy        599 GAGGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGG 658
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1283 GAGGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGG 1342

Qy        659 GCTTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTG 718
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1343 GCTTTCTTTCTGGGTTTGGGCCATTTCAGTTCTTATGTGTGTACTATTCTATCATTATTG 1402

Qy        719 TATAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGA 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1403 TATAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGA 1462

Qy        779 ATAGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAAC 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1463 ATAGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAAC 1522

Qy        839 CCAAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAG 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1523 CCAAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAG 1582

Qy        899 TGTTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTT 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1583 TGTTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTT 1642

Qy        959 CTTTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAA 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1643 CTTTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAA 1702

Qy       1019 TGAAAAGTT 1027
              |||||||||
Db       1703 TGAAAAGTT 1711


RESULT 2
BC017861
LOCUS       BC017861               1510 bp    mRNA    linear   HTC 06-DEC-2001
DEFINITION  Homo sapiens, Similar to T cell receptor gamma locus, clone
            IMAGE:4687960, mRNA.
ACCESSION   BC017861
VERSION     BC017861.1  GI:17389678
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1510)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-DEC-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 36 Row: k Column: 19
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1. .1510
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4687960"
                     /tissue_type="Lung"
                     /clone_lib="NIH_MGC_77"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
ORIGIN

  Query Match             85.6%;  Score 879.2;  DB 3;  Length 1510;
  Best Local Similarity   93.5%;  Pred. No. 2.9e-230;
  Matches  948;  Conservative    0;  Mismatches   18;  Indels   48;  Gaps    1;

Qy         59 CAGATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTG 118
              | |||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db        464 CTGATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCGATTG 523

Qy        119 CTGAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTG 178
              |||||||||| |||||||||||||||||||||||||||||||||||||||||||||| |
Db        524 CTGAAACAAAACTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCAG 583

Qy        179 ATGTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGG 238
              || |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        584 ATATTATTAAGATACATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGG 643

Qy        239 GGAACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAA 298
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 15, 2005, 15:44:35 ; Search time 598 Seconds
(without alignments)
10166.512 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1027
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaaatgaaaagtt 1027

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1025.4	99.8	1027	4	AAF56401	Aaf56401 Human TCR
	2	1025.4	99.8	1027	10	ADB75580	Adb75580 Prostate
	3	1025.4	99.8	1027	13	ADP54862	Adp54862 Human PRO
	4	1012.2	98.6	1395	12	ADL06443	Adl06443 Human tum
	5	966.6	94.1	1799	6	ABZ35381	Abz35381 Human gen
	6	944	91.9	1418	6	AAD38828	Aad38828 Human PSN
c	7	913.4	88.9	2658	10	ADE08798	Ade08798 Novel DNA
	8	906	88.2	1155	3	AAF15623	Aaf15623 Human pro
	9	904	88.0	1560	10	ADE06987	Ade06987 Novel cod
	10	900.6	87.7	1421	6	ABL65007	Abl65007 Lung canc
	11	900.6	87.7	1421	13	ADR65880	Adr65880 Human pro
	12	900.6	87.7	1421	13	ADR66783	Adr66783 Human pro
	13	836.4	81.4	1586	6	ABK64529	Abk64529 Human ben
	14	836.4	81.4	1586	13	ADP54939	Adp54939 Human PRO
	15	833.6	81.2	2658	10	ADE08798	Ade08798 Novel DNA
	16	820	79.8	1046	12	ADP10449	Adp10449 Reference
c	17	795.8	77.5	1162	5	ABV27727	Abv27727 Human pro
c	18	795.8	77.5	1162	5	ABV24153	Abv24153 Human pro
c	19	795.8	77.5	1162	5	ABV24540	Abv24540 Human pro
c	20	795.8	77.5	1162	5	ABV25162	Abv25162 Human pro
c	21	795.8	77.5	1162	5	ABV21526	Abv21526 Human pro
c	22	795.8	77.5	1162	5	ABV27345	Abv27345 Human pro
c	23	795.8	77.5	1162	5	ABV24544	Abv24544 Human pro
c	24	795.8	77.5	1162	5	ABV24851	Abv24851 Human pro
c	25	619	60.3	924	5	ABV25092	Abv25092 Human pro
	26	601	58.5	786	13	ADP55163	Adp55163 Human PRO
	27	521.2	50.7	1080	2	AAQ37617	Aaq37617 TCR gamma
	28	514.8	50.1	1080	1	AAN91698	Aan91698 Sequence
c	29	497.2	48.4	1316	5	ABV26485	Abv26485 Human pro
c	30	497.2	48.4	1316	5	ABV25093	Abv25093 Human pro
c	31	471.6	45.9	486	5	ABV25013	Abv25013 Human pro
	32	450.8	43.9	825	2	AAQ66888	Aaq66888 Human lym
	33	450.8	43.9	825	9	ACF35988	Acf35988 Human T c
	34	446.8	43.5	784	5	ABV21896	Abv21896 Human pro
c	35	444	43.2	511	2	AAV58549	Aav58549 Prostate
c	36	444	43.2	511	2	AAV61291	Aav61291 cDNA sequ
c	37	444	43.2	511	3	AAA06312	Aaa06312 Human imm
c	38	444	43.2	511	3	ABS71217	Abs71217 Human pro
c	39	444	43.2	511	4	AAH93428	Aah93428 Human pro
c	40	444	43.2	511	4	AAS63520	Aas63520 Human pro
c	41	444	43.2	511	4	AAH02493	Aah02493 Prostate
c	42	444	43.2	511	4	AAH84742	Aah84742 Human pro
c	43	444	43.2	511	5	ACA59329	Aca59329 Prostate
c	44	444	43.2	511	5	AAS10071	Aas10071 Human pro
c	45	444	43.2	511	6	ABL94892	Abl94892 Human 1A-

ALIGNMENTS

RESULT 1
AAF56401
ID AAF56401 standard; cDNA; 1027 BP.
XX
AC AAF56401;
XX
DT 12-APR-2001 (first entry)
XX
DE Human TCRgamma alternate reading frame protein coding sequence.
XX
KW Human; TARP; prostate cancer; breast cancer; immunotherapy;
KW T cell receptor gamma alternate reading frame protein; TCRgamma; ss.
XX
OS Homo sapiens.
XX
PN WO200104309-A1.
XX
PD 18-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US019039.
XX
PR 13-JUL-1999; 99US-0143560P.
PR 01-OCT-1999; 99US-0157471P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Pastan I, Essand M, Lee B, Vasmatzis G, Wolfgang C;
XX
DR WPI; 2001-081050/09.
DR P-PSDB; AAB66399, AAB66400.
XX
PT Isolated T-cell receptor gamma alternate reading frame protein useful for
PT diagnosing and raising an immune response to prostate cancer and breast
PT cancer.
XX
PS Example 1; Fig 1; 85pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human T cell receptor alternate reading frame protein (TARP). This
CC protein is expressed in prostate and breast cancer cells at higher levels
CC than normal and so can be used in the immunotherapy of these cancers, as
CC well as their detection and prevention
XX

```
SQ   Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;

  Query Match           99.8%;  Score 1025.4;  DB 4;  Length 1027;
  Best Local Similarity   99.9%;  Pred. No. 2.5e-290;
  Matches 1026;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60

Qy         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120

Qy        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180

Qy        181 GTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTATTAAGATACATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240

Qy        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300

Qy        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360

Qy        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420

Qy        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480

Qy        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540

Qy        541 CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGA 600

Qy        601 GGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660

Qy        661 TTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTA 720

Qy        721 TAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT 780

Qy        781 AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCC 840

Qy        841 AAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTG 900

Qy        901 TTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960

Qy        961 TTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAATG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTGGTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAATG 1020

Qy       1021 AAAAGTT 1027
              |||||||
Db       1021 AAAAGTT 1027


RESULT 2
ADB75580
ID   ADB75580 standard; cDNA; 1027 BP.
XX
AC   ADB75580;
XX
DT   04-DEC-2003  (first entry)
XX
DE   Prostate cancer marker cDNA.
XX
KW   Prostate; cancer; cytostatic; gene therapy; marker; ss.
XX
OS   Homo sapiens.
XX
PN   WO2003009814-A2.
XX
PD   06-FEB-2003.
XX
PF   25-JUL-2002; 2002WO-US023913.
XX
PR   25-JUL-2001; 2001US-0307982P.
PR   22-AUG-2001; 2001US-0314356P.
PR   25-SEP-2001; 2001US-0325020P.
PR   12-DEC-2001; 2001US-0341746P.
PR   05-MAR-2002; 2002US-0362158P.
XX
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbatcheva B;
PI   Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
XX
DR   WPI; 2003-248033/24.
XX
PT   New nucleic acid molecule, useful for diagnosing or treating prostate
PT   cancer.
XX
PS   Disclosure; SEQ ID NO 404; 99pp; English.
XX
CC   The invention relates to newly discovered cancer markers associated with
CC   the cancerous state of prostate cells. Also disclosed is a method of
CC   assessing whether a patient is afflicted with prostate cancer. The method
CC   of the invention involves assessing whether a patient is afflicted with
CC   prostate cancer by comparing the level of expression of a marker in a
CC   patient sample and the normal level of expression of the marker in a
CC   control non-prostate cancer sample, where a significant increase in the
CC   level of expression of the marker in the patient sample and the normal
CC   level indicates that the patient is afflicted with prostate cancer.
CC   Nucleic acids of the invention are useful for diagnosing or treating
CC   prostate cancer, and may be useful in gene therapy. Sequences given in
CC   ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC   data for this patent did not form part of the printed specification, but
CC   was obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;

  Query Match           99.8%;  Score 1025.4;  DB 10;  Length 1027;
  Best Local Similarity   99.9%;  Pred. No. 2.5e-290;
  Matches 1026;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60

Qy         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: October 15, 2005, 15:46:50 ; Search time 4554 Seconds
(without alignments)
10927.428 Million cell updates/sec

Title: US-10-031-158-13
Perfect score: 1027
Sequence: 1 gggcaagagttgggcaaaaa.....atttaaaaaaatgaaaagtt 1027

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1027	100.0	1027	6	AX074415	AX074415 Sequence
2	1025.4	99.8	1027	9	AF151103	AF151103 Homo sapi
3	1011.6	98.5	1825	9	BC072396	BC072396 Homo sapi
4	1007.2	98.1	1054	9	BC062761	BC062761 Homo sapi
5	982.4	95.7	1530	9	HUMTCRGA	M16768 Human T-cel
6	971.4	94.6	1470	9	BC072387	BC072387 Homo sapi
7	939.6	91.5	958	9	HUMTCRGAAD	M27334 Homo sapien
8	902.2	87.8	1655	9	BC039725	BC039725 Homo sapi
9	900.6	87.7	1421	6	AX332835	AX332835 Sequence
10	900.6	87.7	1421	9	HUMTCGXH	M13231 Human T-cel
11	875.8	85.3	1013	9	HUMTCRGAAA	M27331 Homo sapien
12	863.6	84.1	1402	9	HUMTCRGB	M16804 Human T-cel
13	855.8	83.3	1003	9	HUMTCRGAAB	M27332 Homo sapien
14	836.4	81.4	1586	9	HUMTCRGAD	M30894 Human T-cel
15	820	79.8	1046	9	HUMTCGCH	M17323 Homo sapien
16	820	79.8	1046	11	G31759	G31759 sWSS1280 Er
c 17	795.8	77.5	1162	6	CQ489650	CQ489650 Sequence
c 18	795.8	77.5	1162	6	CQ492275	CQ492275 Sequence
c 19	795.8	77.5	1162	6	CQ492662	CQ492662 Sequence
c 20	795.8	77.5	1162	6	CQ492666	CQ492666 Sequence
c 21	795.8	77.5	1162	6	CQ492973	CQ492973 Sequence
c 22	795.8	77.5	1162	6	CQ493284	CQ493284 Sequence
c 23	795.8	77.5	1162	6	CQ495492	CQ495492 Sequence
c 24	795.8	77.5	1162	6	CQ495873	CQ495873 Sequence
25	731.8	71.3	757	9	HUMTCGCJ	M17325 Homo sapien
c 26	619	60.3	924	6	CQ493214	CQ493214 Sequence
27	601	58.5	786	9	HUMTCRGAAC	M27333 Homo sapien
28	589	57.4	847	9	HUMTCGCD3	M14998 Human T-cel
29	589	57.4	140691	9	AF159056	AF159056 Homo sapi
c 30	589	57.4	171816	9	AC006033	AC006033 Homo sapi
31	579.6	56.4	586	9	HUMTCGXO	M15007 Human T-cel
32	576.6	56.1	1041	9	AY190025	AY190025 Macaca mu
c 33	576.4	56.1	66558	2	AC130306	AC130306 Homo sapi
34	575.6	56.0	1041	9	AY190027	AY190027 Macaca mu
35	521.2	50.7	1080	9	HSTCRGR	Y00790 Human mRNA
36	514.8	50.1	1080	6	AR364341	AR364341 Sequence
c 37	497.2	48.4	1316	6	CQ493215	CQ493215 Sequence
c 38	497.2	48.4	1316	6	CQ494607	CQ494607 Sequence
c 39	471.6	45.9	486	6	CQ493135	CQ493135 Sequence
40	471.6	45.9	600	9	HUMTCGCI	M17324 Homo sapien
41	450.8	43.9	825	6	I89883	I89883 Sequence 18
42	450.8	43.9	825	9	HSTCRGSA	X72500 H.sapiens m
43	446.8	43.5	784	6	CQ490020	CQ490020 Sequence
c 44	444	43.2	511	6	BD241985	BD241985 Compounds
c 45	444	43.2	511	6	AR237168	AR237168 Sequence

ALIGNMENTS

RESULT 1
AX074415
LOCUS AX074415 1027 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 13 from Patent WO0104309.
ACCESSION AX074415
VERSION AX074415.1 GI:12710553
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Pastan,I., Essand,M., Lee,B., Vasmatzis,G. and Wolfgang,C.
TITLE T-cell receptor _g(g) alternate reading frame protein, (tarp) and uses thereof
JOURNAL Patent: WO 0104309-A 13 18-JAN-2001;
UNITED STATES GOVERNMENT (US)
FEATURES Location/Qualifiers
source 1. .1027
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS 74. .250
/note="unnamed protein product; Coding region for PS-TCR gamma 1 polypeptide (TARP)"
/codon_start=1
/protein_id="CAC28473.1"
/db_xref="GI:12710554"
/translation="MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP"
CDS 247. .582
/note="unnamed protein product; Coding region for PS-TCR gamma 2 polypeptide (deduced amino acid sequence not displayed along with DNA sequence, due to overlapping CDS's)"
/codon_start=1
/protein_id="CAC28474.1"
/db_xref="GI:12710555"
/translation="MKTNDTYMKFSWLTVPEKSLDKEHRCIVRHENNKNGVDQEIIFPPIKTDVITMDPKDNCSKDANDTLLLQLTNTSAYYMYLLLLLKSVVYFAIITCCLLRRTAFCCNGEKS"
ORIGIN

```
  Query Match             100.0%;  Score 1027;  DB 6;  Length 1027;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-261;
  Matches 1027;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAAGCTTATCATTACA 60

Qy         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCT 120

Qy        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GAAACAAAGCTCCAGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGAT 180

Qy        181 GTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GTTATTAAGATACATTGGGAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGG 240

Qy        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACACCATGAAGACTAACGACACATACATGAAATTTAGCTGGTTAACGGTGCCAGAAAAG 300

Qy        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCACTGGACAAAGAACACAGATGTATCGTCAGACATGAGAATAATAAAAACGGAGTTGAT 360

Qy        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CAAGAAATTATCTTTCCTCCAATAAAGACGGATGTCATCACAATGGATCCCAAAGACAAT 420

Qy        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGTTCAAAAGATGCAAATGATACACTACTGCTGCAGCTCACAAACACCTCTGCATATTAC 480

Qy        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGTACCTCCTCCTGCTCCTCAAGAGTGTGGTCTATTTTGCCATCATCACCTGCTGTCTG 540

Qy        541 CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CTTAGAAGAACGGCTTTCTGCTGCAATGGAGAGAAATCATAACAGACGGTGGCACAAGGA 600

Qy        601 GGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCCATCTTTTCCTCATCGGTTATTGTCCCTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660

Qy        661 TTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTA 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTTCTTTCTGGGTTTGGGCCATTTCAGTTCTCATGTGTGTACTATTCTATCATTATTGTA 720

Qy        721 TAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TAACGGTTTTCAAACCAGTGGGCACACAGAGAACCTCACTCTGTAATAACAATGAGGAAT 780

Qy        781 AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AGCCACGGCGATCTCCAGCACCAATCTCTCCATGTTTTCCACAGCTCCTCCAGCCAACCC 840

Qy        841 AAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AAATAGCGCCTGCTATAGTGTAGACATCCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTG 900

Qy        901 TTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTCTTTAATCAGATAACTGCCTGGAAGCCTTTCATTTTACACGCCCTGAAGCAGTCTTCT 960

Qy        961 TTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAATG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTGCTAGTTGAATTATGTGGTGTGTTTTTCCGTAATAAGCAAAATAAATTTAAAAAAATG 1020
```

```
Qy       1021 AAAAGTT 1027
              |||||||
Db       1021 AAAAGTT 1027
```

```
RESULT 2
AF151103
LOCUS       AF151103                1027 bp    mRNA    linear   PRI 03-NOV-2000
DEFINITION  Homo sapiens TCRgamma alternate reading frame protein (TCRg) mRNA,
            complete cds.
ACCESSION   AF151103
VERSION     AF151103.1  GI:5758136
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Essand,M., Vasmatzis,G., Brinkmann,U., Duray,P., Lee,B. and
            Pastan,I.
  TITLE     High expression of a specific T-cell receptor gamma transcript in
            epithelial cells of the prostate
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (16), 9287-9292 (1999)
  MEDLINE   99362750
   PUBMED   10430935
REFERENCE   2  (bases 1 to 1027)
  AUTHORS   Wolfgang,C.D., Essand,M., Vincent,J.J., Lee,B. and Pastan,I.
  TITLE     TARP: a nuclear protein expressed in prostate and breast cancer
            cells derived from an alternate reading frame of the T cell
            receptor gamma chain locus
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9437-9442 (2000)
  MEDLINE   20402552
   PUBMED   10931945
REFERENCE   3  (bases 1 to 1027)
  AUTHORS   Essand,M., Vasmatzis,G., Brinkmann,U., Duray,P., Lee,B. and
            Pastan,I.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-1999) Division of Basic Sciences, Laboratory of
            Molecular Biology, National Institutes of Health, National Cancer
            Institute, 37 Convent Drive, Bethesda, MD 20892, USA
FEATURES             Location/Qualifiers
     source          1. .1027
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /note="transcript from an unrearranged TCRg locus"
     gene            1. .1027
                     /gene="TCRg"
     J_segment       1. .60
                     /gene="TCRg"
                     /note="gamma 1.2"
     C_region        61. .390
                     /gene="TCRg"
                     /note="gamma 1; corresponds to exon I"
     CDS             74. .250
                     /gene="TCRg"
                     /note="TARP"
                     /codon_start=1
                     /product="TCRgamma alternate reading frame protein"
                     /protein_id="AAG29337.1"
                     /db_xref="GI:11093514"
                     /translation="MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKK
                     RRATRFWDPRRGTP"
     C_region        391. .438
                     /gene="TCRg"
                     /note="gamma 1; corresponds to exon II"
     C_region        439. .579
                     /gene="TCRg"
                     /note="gamma 1; corresponds to exon III"
ORIGIN
```

```
  Query Match             99.8%;  Score 1025.4;  DB 9;  Length 1027;
  Best Local Similarity   99.9%;  Pred. No. 1.5e-260;
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 06:48:32 ; Search time 707 Seconds
(without alignments)
575.070 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFPFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 8765912 seqs, 3504951483 residues

Total number of hits satisfying chosen parameters: 17531824

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135741_18954/app_query.fasta_1.199
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10031158_@CGN_1_1_723_@runat_14102005_135741_18954
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=10C
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		307	100.0	470	9	US-09-864-761-15428	Sequence 15428, A
2		307	100.0	477	9	US-09-864-761-15438	Sequence 15438, A
c 3		307	100.0	539	20	US-10-357-930-22784	Sequence 22784, A
c 4		307	100.0	539	20	US-10-357-930-28636	Sequence 28636, A
5		307	100.0	825	16	US-10-340-536-29	Sequence 29, Appl
6		307	100.0	1027	15	US-10-205-823-404	Sequence 404, App
7		307	100.0	1027	22	US-10-756-149-2859	Sequence 2859, Ap
8		307	100.0	1027	24	US-11-051-454-404	Sequence 404, App
c 9		307	100.0	1316	20	US-10-357-930-25082	Sequence 25082, A
c 10		307	100.0	1316	20	US-10-357-930-26474	Sequence 26474, A
11		307	100.0	1395	20	US-10-643-795A-23	Sequence 23, Appl
12		307	100.0	1395	21	US-10-948-518-23	Sequence 23, Appl
13		307	100.0	1418	10	US-09-957-708-16	Sequence 16, Appl
14		307	100.0	124933	22	US-10-737-082-97	Sequence 97, Appl
15		307	100.0	124933	22	US-10-765-790-97	Sequence 97, Appl
c 16		304	99.0	486	20	US-10-357-930-25002	Sequence 25002, A
17		302	98.4	316	9	US-09-864-761-31943	Sequence 31943, A
c 18		298	97.1	569	20	US-10-357-930-43494	Sequence 43494, A
19		294	95.8	1421	9	US-09-954-456-317	Sequence 317, App
20		294	95.8	1421	21	US-10-843-641A-3344	Sequence 3344, Ap
21		294	95.8	1799	15	US-10-101-510-492	Sequence 492, App
22		284	92.5	1586	10	US-09-960-706-676	Sequence 676, App
23		284	92.5	1586	10	US-09-873-319-424	Sequence 424, App
24		284	92.5	1586	21	US-10-847-918-14	Sequence 14, Appl
25		259	84.4	533	20	US-10-357-930-37875	Sequence 37875, A
26		245	79.8	1155	9	US-09-925-300-58	Sequence 58, Appl
c 27		240	78.2	1022	20	US-10-357-930-23803	Sequence 23803, A
c 28		240	78.2	1022	20	US-10-357-930-23975	Sequence 23975, A
c 29		240	78.2	1022	20	US-10-357-930-23980	Sequence 23980, A
c 30		240	78.2	1022	20	US-10-357-930-24023	Sequence 24023, A
c 31		240	78.2	1022	20	US-10-357-930-24061	Sequence 24061, A
c 32		240	78.2	1022	20	US-10-357-930-24062	Sequence 24062, A
c 33		240	78.2	1022	20	US-10-357-930-24097	Sequence 24097, A
c 34		240	78.2	1022	20	US-10-357-930-24100	Sequence 24100, A
c 35		240	78.2	1022	20	US-10-357-930-24113	Sequence 24113, A
c 36		240	78.2	1022	20	US-10-357-930-25188	Sequence 25188, A
c 37		240	78.2	1022	20	US-10-357-930-25812	Sequence 25812, A
c 38		240	78.2	1022	20	US-10-357-930-29708	Sequence 29708, A
c 39		240	78.2	1022	20	US-10-357-930-29892	Sequence 29892, A
c 40		240	78.2	1022	20	US-10-357-930-29936	Sequence 29936, A
c 41		240	78.2	1022	20	US-10-357-930-29971	Sequence 29971, A
c 42		240	78.2	1022	20	US-10-357-930-29972	Sequence 29972, A
c 43		240	78.2	1022	20	US-10-357-930-30007	Sequence 30007, A
c 44		227	73.9	475	20	US-10-357-930-39216	Sequence 39216, A
c 45		216	70.4	401	20	US-10-357-930-13495	Sequence 13495, A

ALIGNMENTS

RESULT 1
US-09-864-761-15428
; Sequence 15428, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

```
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15428
;   LENGTH: 470
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF159056.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
US-09-864-761-15428

Alignment Scores:
Pred. No.:              9.57e-36        Length:         470
Score:                  307.00          Matches:        58
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-10-031-158-14 (1-58) x US-09-864-761-15428 (1-470)

Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 262

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 322

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 376

RESULT 2
US-09-864-761-15438
; Sequence 15438, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15438
;   LENGTH: 477
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF159056.1
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
US-09-864-761-15438

Alignment Scores:
Pred. No.:              9.76e-36        Length:         477
Score:                  307.00          Matches:        58
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-10-031-158-14 (1-58) x US-09-864-761-15438 (1-477)

Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        212 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 271

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 06:07:46 ; Search time 3111 Seconds
(without alignments)
709.652 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135739_18853/app_query.fasta_1.199
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158_@CGN_1_1_5180_@runat_14102005_135739_18853 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	307	100.0	427	2	BF679325	BF679325 602153825
2	307	100.0	510	1	AV716204	AV716204 AV716204
3	307	100.0	596	1	AV715641	AV715641 AV715641
4	307	100.0	603	1	AV714189	AV714189 AV714189
5	307	100.0	803	1	AV756492	AV756492 AV756492
6	307	100.0	818	4	BI771070	BI771070 603059635
7	307	100.0	821	2	BF679123	BF679123 602153390
8	307	100.0	861	2	BF681385	BF681385 602156579
9	307	100.0	865	1	AI557112	AI557112 PT2.1_13_
10	307	100.0	881	2	BF677648	BF677648 602085529
11	307	100.0	902	2	BF674593	BF674593 602137811
12	307	100.0	927	2	BF679476	BF679476 602153280
13	307	100.0	1058	4	BM544213	BM544213 AGENCOURT
14	307	100.0	3533	3	BC030554	BC030554 Homo sapi
15	304	99.0	789	5	BU199143	BU199143 DCBCIA08
16	294	95.8	653	2	BE974171	BE974171 601680522
17	294	95.8	1005	5	BM919380	BM919380 AGENCOURT
18	294	95.8	1757	3	BC039116	BC039116 Homo sapi
c 19	290	94.5	434	2	BF747659	BF747659 QV2-BT061
c 20	290	94.5	497	2	AW087486	AW087486 xb67c01.x
21	290	94.5	681	1	AV714756	AV714756 AV714756
22	290	94.5	1510	3	BC017861	BC017861 Homo sapi
23	289	94.1	1055	4	BF964904	BF964904 602268596
24	288	93.8	523	2	BF675605	BF675605 602083495
c 25	284	92.5	362	1	AI758378	AI758378 ty68c12.x
26	278	90.6	719	2	BF681238	BF681238 602155502
27	270	87.9	854	2	BF681029	BF681029 602155433
28	268	87.3	629	2	BF679165	BF679165 602153737
29	267	87.0	803	2	BF680840	BF680840 602156088
30	262	85.3	772	4	BI836020	BI836020 603084004
31	259	84.4	808	2	BF673289	BF673289 602136139
32	253	82.4	857	2	BF674457	BF674457 602137231
33	251	81.8	849	2	BF679284	BF679284 602153477
34	248	80.8	995	7	CO647648	CO647648 ILLUMIGEN
35	245	79.8	663	6	CB553506	CB553506 MMSP0044_
36	227	73.9	870	2	BF678129	BF678129 602085181
37	226	73.6	637	6	CB551454	CB551454 MMSP0013_
38	224	73.0	670	2	BF677916	BF677916 602084736
39	221	72.0	827	4	BG535408	BG535408 602563069
40	209	68.1	572	2	BF370026	BF370026 MR3-FN000
41	204.5	66.6	864	4	BI834993	BI834993 603088885
42	203	66.1	552	5	BP164784	BP164784 BP164784
43	200	65.1	572	5	BP162849	BP162849 BP162849
44	200	65.1	888	5	BP162346	BP162346 BP162346
45	194	63.2	744	7	CR791490	CR791490 DKFZp468F

ALIGNMENTS

RESULT 1
BF679325
LOCUS BF679325 427 bp mRNA linear EST 21-DEC-2000
DEFINITION 602153825F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4294566 5', mRNA sequence.
ACCESSION BF679325
VERSION BF679325.1 GI:11953220
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 427)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1144 row: b column: 07
High quality sequence stop: 427.
FEATURES Location/Qualifiers
source 1. .427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4294566"

/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Alignment Scores:

Pred. No.:	1.22e-30	Length:	427
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679325 (1-427)

```
Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 125

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 185

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 239
```

RESULT 2
AV716204

LOCUS	AV716204	510 bp	mRNA	linear	EST 11-OCT-2000

DEFINITION AV716204 DCB Homo sapiens cDNA clone DCBAMG08 5', mRNA sequence.
ACCESSION AV716204
VERSION AV716204.1 GI:10797721
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 510)
AUTHORS Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X., Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Cheng,Z. and Han,Z.
TITLE Homo sapiens cDNA DCB clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES Location/Qualifiers
source 1. .510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBAMG08"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"
/note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

ORIGIN

Alignment Scores:

Pred. No.:	1.51e-30	Length:	510
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x AV716204 (1-510)

```
Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        159 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 218

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 278

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 332
```

RESULT 3
AV715641

LOCUS	AV715641	596 bp	mRNA	linear	EST 11-OCT-2000

DEFINITION AV715641 DCB Homo sapiens cDNA clone DCBBSB05 5', mRNA sequence.
ACCESSION AV715641
VERSION AV715641.1 GI:10797158
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 596)
AUTHORS Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X., Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Cheng,Z. and Han,Z.
TITLE Homo sapiens cDNA DCB clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES Location/Qualifiers
source 1. .596
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBBSB05"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"
/note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"

ORIGIN

Alignment Scores:

Pred. No.:	1.82e-30	Length:	596
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x AV715641 (1-596)

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 06:00:36 ; Search time 432 Seconds
(without alignments)
794.780 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135738_18834/app_query.fasta_1.199
-DB=N_Geneseq_16Dec04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158_@CGN_1_1_708_@runat_14102005_135738_18834 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	307	100.0	426	3	AAC00681	Aac00681 Human sec
2	307	100.0	470	4	AAK26982	Aak26982 Human bon
3	307	100.0	470	6	ABS01569	Abs01569 Human gen
4	307	100.0	477	4	ABA43588	Aba43588 Human bre
5	307	100.0	477	4	AAK27758	Aak27758 Human bon
6	307	100.0	477	6	ABS02210	Abs02210 Human gen
c 7	307	100.0	539	5	ABV28618	Abv28618 Human pro
c 8	307	100.0	539	5	ABV22791	Abv22791 Human pro
9	307	100.0	825	2	AAQ66888	Aaq66888 Human lym
10	307	100.0	825	9	ACF35988	Acf35988 Human T c
11	307	100.0	1027	4	AAF56401	Aaf56401 Human TCR
12	307	100.0	1027	10	ADB75580	Adb75580 Prostate
13	307	100.0	1027	13	ADP54862	Adp54862 Human PRO
c 14	307	100.0	1316	5	ABV26485	Abv26485 Human pro
c 15	307	100.0	1316	5	ABV25093	Abv25093 Human pro
16	307	100.0	1395	12	ADL06443	Adl06443 Human tum
17	307	100.0	1418	6	AAD38828	Aad38828 Human PSN
c 18	307	100.0	2658	10	ADE08798	Ade08798 Novel DNA
c 19	304	99.0	486	5	ABV25013	Abv25013 Human pro
20	302	98.4	316	4	ABA47985	Aba47985 Human bre
21	302	98.4	316	4	AAK40015	Aak40015 Human bon
22	302	98.4	316	6	ABS14101	Abs14101 Human gen
c 23	298	97.1	569	5	ABV43475	Abv43475 Human pro
24	294	95.8	1421	6	ABL65007	Abl65007 Lung canc
25	294	95.8	1421	13	ADR65880	Adr65880 Human pro
26	294	95.8	1421	13	ADR66783	Adr66783 Human pro
27	294	95.8	1560	10	ADE06987	Ade06987 Novel cod
28	294	95.8	1799	6	ABZ35381	Abz35381 Human gen
29	290	94.5	1080	2	AAQ37617	Aaq37617 TCR gamma
30	290	94.5	2658	10	ADE08798	Ade08798 Novel DNA
31	284	92.5	1046	12	ADP10449	Adp10449 Reference
32	284	92.5	1586	6	ABK64529	Abk64529 Human ben
33	284	92.5	1586	13	ADP54939	Adp54939 Human PRO
34	279	90.9	1080	1	AAN91698	Aan91698 Sequence
35	264	86.0	786	13	ADP55163	Adp55163 Human PRO
36	259	84.4	533	5	ABV37856	Abv37856 Human pro
37	251	81.8	460	3	AAC04122	Aac04122 Human sec
38	245	79.8	1155	3	AAF15623	Aaf15623 Human pro
c 39	240	78.2	1022	5	ABV29989	Abv29989 Human pro
c 40	240	78.2	1022	5	ABV24072	Abv24072 Human pro
c 41	240	78.2	1022	5	ABV24124	Abv24124 Human pro
c 42	240	78.2	1022	5	ABV25823	Abv25823 Human pro
c 43	240	78.2	1022	5	ABV24034	Abv24034 Human pro
c 44	240	78.2	1022	5	ABV23986	Abv23986 Human pro
c 45	240	78.2	1022	5	ABV23991	Abv23991 Human pro

ALIGNMENTS

RESULT 1
AAC00681
ID AAC00681 standard; cDNA; 426 BP.
XX
AC AAC00681;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 679.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG00675.

```
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
PS   Claim 1; SEQ ID NO 679; 71pp + Sequence Listing; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. An ORF has been identified within the
CC   sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors
XX
SQ   Sequence 426 BP; 151 A; 83 C; 88 G; 104 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                5.17e-34        Length:         426
Score:                    307.00          Matches:        58
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       3               Gaps:           0

US-10-031-158-14 (1-58) x AAC00681 (1-426)

Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 153

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 213

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 267

RESULT 2
AAK26982
ID   AAK26982 standard; DNA; 470 BP.
XX
AC   AAK26982;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 1539.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000668.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human bone marrow.
XX
PS   Example 4; SEQ ID NO 1539; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention
XX
SQ   Sequence 470 BP; 148 A; 102 C; 87 G; 133 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:                5.9e-34         Length:         470
Score:                    307.00          Matches:        58
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       4               Gaps:           0

US-10-031-158-14 (1-58) x AAK26982 (1-470)

Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 262

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        263 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 322

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 376

RESULT 3
ABS01569
ID   ABS01569 standard; DNA; 470 BP.
XX
AC   ABS01569;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe from lung SEQ ID No 1560.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
```

GenCore version 5.1.6

OM protein - nucleic search, using frame_plus_p2n model

Run on: October 16, 2005, 06:01:41 ; Search time 1790 Seconds
(without alignments)
1570.057 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10031158/runat_14102005_135739_18841/app_query.fasta_1.199
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10031158_@CGN_1_1_5600_@runat_14102005_135739_18841 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	307	100.0	410	9	HUMTCGCD1	M14996 Human T-cel
2	307	100.0	426	6	AX884816	AX884816 Sequence
3	307	100.0	426	6	BD024426	BD024426 Sequence
4	307	100.0	470	6	CQ131517	CQ131517 Sequence
5	307	100.0	470	6	CQ290455	CQ290455 Sequence
6	307	100.0	477	6	CQ132293	CQ132293 Sequence
7	307	100.0	477	6	CQ200006	CQ200006 Sequence
8	307	100.0	477	6	CQ291096	CQ291096 Sequence
c 9	307	100.0	539	6	CQ490917	CQ490917 Sequence
c 10	307	100.0	539	6	CQ496769	CQ496769 Sequence
11	307	100.0	723	9	AJ583012	AJ583012 Homo sapi
12	307	100.0	726	9	AJ583014	AJ583014 Homo sapi
13	307	100.0	825	6	I89883	I89883 Sequence 18
14	307	100.0	825	9	HSTCRGSA	X72500 H.sapiens m
15	307	100.0	958	9	HUMTCRGAAD	M27334 Homo sapien
16	307	100.0	1027	6	AX074415	AX074415 Sequence
17	307	100.0	1027	9	AF151103	AF151103 Homo sapi
18	307	100.0	1054	9	BC062761	BC062761 Homo sapi
c 19	307	100.0	1316	6	CQ493215	CQ493215 Sequence
c 20	307	100.0	1316	6	CQ494607	CQ494607 Sequence
21	307	100.0	1470	9	BC072387	BC072387 Homo sapi
22	307	100.0	1530	9	HUMTCRGA	M16768 Human T-cel
23	307	100.0	1825	9	BC072396	BC072396 Homo sapi
24	307	100.0	66558	2	AC130306	AC130306 Homo sapi
25	307	100.0	140691	9	AF159056	AF159056 Homo sapi
c 26	307	100.0	171816	9	AC006033	AC006033 Homo sapi
27	304	99.0	400	9	HSTCELG1	X02766 Human gene
c 28	304	99.0	486	6	CQ493135	CQ493135 Sequence
29	302	98.4	316	6	CQ144550	CQ144550 Sequence
30	302	98.4	316	6	CQ204403	CQ204403 Sequence
31	302	98.4	316	6	CQ302987	CQ302987 Sequence
c 32	298	97.1	569	6	CQ511627	CQ511627 Sequence
33	294	95.8	330	9	HUMTCGCG	M15002 Human T-cel
34	294	95.8	635	9	HSTRGC54	X06776 Human trunc
35	294	95.8	720	9	HSTRGC10	X06775 Human aberr
36	294	95.8	1013	9	HUMTCRGAAA	M27331 Homo sapien
37	294	95.8	1160	9	HSTRGC64	X06774 Human mRNA
38	294	95.8	1402	9	HUMTCRGB	M16804 Human T-cel
39	294	95.8	1421	6	AX332835	AX332835 Sequence
40	294	95.8	1421	9	HUMTCGXH	M13231 Human T-cel
41	294	95.8	1655	9	BC039725	BC039725 Homo sapi
42	292	95.1	757	9	HUMTCGCJ	M17325 Homo sapien
43	290	94.5	201	11	BV205008	BV205008 sqnm21701
44	290	94.5	300	9	HUMTCGCE	M14999 Human T-cel
45	290	94.5	330	9	HUMTCGCF	M15001 Human T-cel

ALIGNMENTS

RESULT 1
HUMTCGCD1
LOCUS HUMTCGCD1 410 bp DNA linear PRI 14-SEP-1995
DEFINITION Human T-cell receptor germline gamma-chain (TCRGC1) gene, C1-region, exon 1.
ACCESSION M14996
VERSION M14996.1 GI:339076
KEYWORDS C-region; T cell receptor gamma-chain; T-cell receptor; germline.
SEGMENT 1 of 3
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 410)
AUTHORS Lefranc,M.P., Forster,A. and Rabbitts,T.H.
TITLE Genetic polymorphism and exon changes of the constant regions of the human T-cell rearranging gene gamma
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (24), 9596-9600 (1986)
MEDLINE 87092294
PUBMED 2879283
COMMENT Original source text: Homo sapiens (clone: lambda-D19) DNA.
Entry revisions for [1] kindly provided by M.-P. Lefranc, 22-JUN-1988.
FEATURES Location/Qualifiers
source 1. .410
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```
                     /map="7p15-p14"
                     /clone="lambda-D19"
                     /cell_line="D-PLL"
     exon            <41. .370
                     /gene="TCRGC1"
                     /note="G00-120-408"
                     /number=1
ORIGIN      205 bp upstream of BamHI site.

Alignment Scores:
Pred. No.:              3.7e-30         Length:         410
Score:                  307.00          Matches:        58
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     9               Gaps:           0

US-10-031-158-14 (1-58) x HUMTCGCD1 (1-410)

Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         53 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 112

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 172

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 226

RESULT 2
AX884816
LOCUS       AX884816                426 bp    DNA     linear   PAT 18-DEC-2003
DEFINITION  Sequence 679 from Patent EP1033401.
ACCESSION   AX884816
VERSION     AX884816.1  GI:40040355
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Dumas Milne Edwards,J.B., Duclert,A. and Giordano,J.Y.
  TITLE     Expressed sequence tags and encoded human proteins
  JOURNAL   Patent: EP 1033401-A 679 06-SEP-2000;
            Genset (FR)
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             267. .>425
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAE98930.1"
                     /db_xref="GI:40040356"
                     /translation="MKTNDTYMKFSWLTVPEKSLDKEHRCIVRHENNKNGVDQEIIFP
                     PIKTDVITM"
ORIGIN

Alignment Scores:
Pred. No.:              3.86e-30        Length:         426
Score:                  307.00          Matches:        58
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-10-031-158-14 (1-58) x AX884816 (1-426)

Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 153

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 213

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 267

RESULT 3
BD024426
LOCUS       BD024426                426 bp    DNA     linear   PAT 27-AUG-2002
DEFINITION  Sequence tag and encoded human protein.
ACCESSION   BD024426
VERSION     BD024426.1  GI:22565649
KEYWORDS    JP 2001269182-A/672.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 426)
  AUTHORS   Edwards,J.B.D.M., Duclair,E. and Jordan,J.Y.
  TITLE     Sequence tag and encoded human protein
  JOURNAL   Patent: JP 2001269182-A 672 02-OCT-2001;
            GENSET
COMMENT     OS   Homo sapiens (human)
            PN   JP 2001269182-A/672
            PD   02-OCT-2001
            PF   24-FEB-2000 JP 2000118773
            PR   26-FEB-1999 US    60/122487
            PI   JEAN BAPUTIST DUMAS MILNE EDWARDS,EIMERIC DUCLAIR,JEAN YVES
            PI   JORDAN
            PC   C12N15/09,C07K14/435,C07K16/18,C12N1/15,C12N1/19,C12N1/21, PC
             C12N5/10,
            PC   C12P21/02,C12P21/08,C12Q1/68//G06F17/30,C12N15/00,C12N5/00, PC
              G06F15/40
            CC
            FH   Key             Location/Qualifiers
            FT   CDS             267. .425.
FEATURES             Location/Qualifiers
     source          1. .426
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              3.86e-30        Length:         426
Score:                  307.00          Matches:        58
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-10-031-158-14 (1-58) x BD024426 (1-426)

Qy          1 MetGlnMetProProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 ATGCAGATGTTTCCCCCAAGCCCACTATTTTTCTTCCTTCAATTGCTGAAACAAAGCTCC 153

Qy         21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 AGAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTTCCCTGATGTTATTAAGATAC 213

Qy         41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        214 ATTGGCAAGAAAAGAAGAGCAACACGATTCTGGGATCCCAGGAGGGGAACACCA 267

RESULT 4
CQ131517
LOCUS       CQ131517                470 bp    DNA     linear   PAT 21-JAN-2004
DEFINITION  Sequence 1539 from Patent WO0157276.
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:43:34 ; Search time 116 Seconds
(without alignments)
208.359 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1859788 seqs, 416717961 residues

Total number of hits satisfying chosen parameters: 1859788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	307	100.0	58	14	US-10-205-823-405	Sequence 405, App
2	307	100.0	58	18	US-10-756-149-5551	Sequence 5551, Ap
3	307	100.0	58	20	US-11-051-454-405	Sequence 405, App
4	71.5	23.3	100	15	US-10-424-599-206785	Sequence 206785,
5	67	21.8	81	16	US-10-425-115-189599	Sequence 189599,
6	64.5	21.0	65	16	US-10-425-115-345446	Sequence 345446,
7	63	20.5	86	16	US-10-437-963-186030	Sequence 186030,
8	63	20.5	934	9	US-09-891-216-12	Sequence 12, Appl
9	63	20.5	934	9	US-09-891-216-15	Sequence 15, Appl
10	63	20.5	934	14	US-10-168-402-2	Sequence 2, Appli
11	63	20.5	934	14	US-10-168-402-39	Sequence 39, Appl
12	62	20.2	83	15	US-10-424-599-176153	Sequence 176153,
13	60.5	19.7	325	15	US-10-425-114-72973	Sequence 72973, A
14	60.5	19.7	572	15	US-10-424-599-280962	Sequence 280962,
15	60	19.5	160	16	US-10-437-963-109049	Sequence 109049,
16	60	19.5	900	15	US-10-282-122A-68765	Sequence 68765, A
17	59	19.2	84	16	US-10-425-115-229997	Sequence 229997,
18	59	19.2	108	16	US-10-437-963-193767	Sequence 193767,
19	59	19.2	641	20	US-11-097-143-5949	Sequence 5949, Ap
20	59	19.2	2176	20	US-11-097-143-10950	Sequence 10950, A
21	58.5	19.1	128	16	US-10-425-115-315195	Sequence 315195,
22	58.5	19.1	440	16	US-10-425-115-307562	Sequence 307562,
23	58.5	19.1	527	15	US-10-425-114-59801	Sequence 59801, A
24	58.5	19.1	551	15	US-10-425-114-65527	Sequence 65527, A
25	58	18.9	83	15	US-10-424-599-256212	Sequence 256212,
26	58	18.9	201	16	US-10-767-701-61188	Sequence 61188, A
27	57.5	18.7	156	16	US-10-437-963-110934	Sequence 110934,
28	57.5	18.7	651	17	US-10-732-923-19734	Sequence 19734, A
29	57.5	18.7	673	15	US-10-369-493-14690	Sequence 14690, A
30	57.5	18.7	673	15	US-10-369-493-15173	Sequence 15173, A
31	57.5	18.7	703	15	US-10-369-493-11508	Sequence 11508, A
32	57.5	18.7	1003	16	US-10-437-963-190584	Sequence 190584,
33	57	18.6	200	16	US-10-425-115-299370	Sequence 299370,
34	56	18.2	78	14	US-10-106-698-5379	Sequence 5379, Ap
35	56	18.2	361	16	US-10-739-930-6604	Sequence 6604, Ap
36	56	18.2	361	17	US-10-732-923-2960	Sequence 2960, Ap
37	56	18.2	459	15	US-10-369-493-19269	Sequence 19269, A
38	55.5	18.1	48	15	US-10-424-599-170289	Sequence 170289,
39	55.5	18.1	66	16	US-10-425-115-189569	Sequence 189569,
40	55.5	18.1	79	15	US-10-424-599-270832	Sequence 270832,
41	55.5	18.1	79	16	US-10-425-115-284110	Sequence 284110,
42	55.5	18.1	99	15	US-10-424-599-240053	Sequence 240053,
43	55.5	18.1	124	15	US-10-424-599-228996	Sequence 228996,
44	55.5	18.1	141	9	US-09-771-161A-99	Sequence 99, Appl
45	55.5	18.1	153	15	US-10-424-599-166111	Sequence 166111,

ALIGNMENTS

RESULT 1
US-10-205-823-405
; Sequence 405, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 405
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-205-823-405

  Query Match             100.0%;  Score 307;  DB 14;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-33;
  Matches   58;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58


RESULT 2
US-10-756-149-5551
; Sequence 5551, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
;  TITLE OF INVENTION:  METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
;  FILE REFERENCE: file
;  CURRENT APPLICATION NUMBER: US/10/756,149
;  CURRENT FILING DATE:  2004-01-12
;  NUMBER OF SEQ ID NOS: 5818
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5551
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-756-149-5551

  Query Match             100.0%;  Score 307;  DB 18;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-33;
  Matches   58;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58


RESULT 3
US-11-051-454-405
; Sequence 405, Application US/11051454
; Publication No. US20050191673A1
; GENERAL INFORMATION:
;  APPLICANT: Schlegel, Robert
;  APPLICANT:  Monahan, John E.
;  APPLICANT:  Endege, Wilson O.
;  APPLICANT:  Gannavarapu, Manjula
;  APPLICANT:  Gorbatcheva, Bella
;  APPLICANT:  Hoersch, Sebastian
;  APPLICANT:  Kamatkar, Shubhangi
;  APPLICANT:  Wonsey, Angela M.
;  APPLICANT:  Glatt, Karen
;  APPLICANT:  Zhao, Xumei
;  APPLICANT:  Anderson, Dustin
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
;  TITLE OF INVENTION:  METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF PROSTATE CANCER
;  FILE REFERENCE: MRI-044
;  CURRENT APPLICATION NUMBER: US/11/051,454
;  CURRENT FILING DATE:  2005-02-04
;  PRIOR APPLICATION NUMBER: US/10/205,823
;  PRIOR FILING DATE: 2002-07-25
;  PRIOR APPLICATION NUMBER: 60/307,982
;  PRIOR FILING DATE: 2001-07-25
;  PRIOR APPLICATION NUMBER: 60/314,356
;  PRIOR FILING DATE: 2001-08-22
;  PRIOR APPLICATION NUMBER: 60/325,020
;  PRIOR FILING DATE: 2001-09-25
;  PRIOR APPLICATION NUMBER: 60/341,746
;  PRIOR FILING DATE: 2001-12-12
;  PRIOR APPLICATION NUMBER: 60/362,158
;  PRIOR FILING DATE: 2002-03-05
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 405
;   LENGTH: 58
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-051-454-405

  Query Match             100.0%;  Score 307;  DB 20;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-33;
  Matches   58;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58


RESULT 4
US-10-424-599-206785
; Sequence 206785, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206785
;   LENGTH: 100
;   TYPE: PRT
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_28754C.1.pep
US-10-424-599-206785

  Query Match             23.3%;  Score 71.5;  DB 15;  Length 100;
  Best Local Similarity   42.0%;  Pred. No. 0.16;
  Matches   21;  Conservative    4;  Mismatches   14;  Indels   11;  Gaps    2;

Qy          1 MQMFPPSPLFFF---LQLLKQSSRRLEHT--------FVFLRNFSLMLLR 39
              :|:| | || ||   |  |   |  || :        | | | |||:|||
Db          9 LQIFSPDPLLFFLSSLSFLDWFSLPLERSDLLFMKSYFSFRRRFSLILLR 58


RESULT 5
US-10-425-115-189599
; Sequence 189599, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa, Thomas J.
;  APPLICANT:  Kovalic, David K.
;  APPLICANT:  Zhou, Yihua
;  APPLICANT:  Cao, Yongwei
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 189599
;   LENGTH: 81
;   TYPE: PRT
;   ORGANISM: Zea mays
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:35:23 ; Search time 55 Seconds
(without alignments)
78.721 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62.5	20.4	77	4	US-09-248-796A-25511	Sequence 25511, A
2	60.5	19.7	72	4	US-09-328-352-7350	Sequence 7350, Ap
3	60	19.5	904	4	US-09-543-681A-4564	Sequence 4564, Ap
4	56	18.2	542	4	US-09-902-540-16320	Sequence 16320, A
5	55.5	18.1	229	4	US-09-489-039A-9391	Sequence 9391, Ap
6	55.5	18.1	316	1	US-08-403-634-4	Sequence 4, Appli
7	55.5	18.1	316	1	US-08-403-634-31	Sequence 31, Appl
8	55.5	18.1	316	3	US-08-913-441B-4	Sequence 4, Appli
9	55.5	18.1	316	3	US-08-913-441B-31	Sequence 31, Appl
10	55.5	18.1	316	4	US-09-571-985C-4	Sequence 4, Appli
11	55.5	18.1	316	4	US-09-571-985C-31	Sequence 31, Appl
12	55	17.9	403	4	US-09-802-213-5	Sequence 5, Appli
13	55	17.9	414	4	US-09-802-213-3	Sequence 3, Appli
14	54	17.6	354	4	US-09-583-110-4854	Sequence 4854, Ap
15	54	17.6	364	4	US-09-107-433-4701	Sequence 4701, Ap
16	53.5	17.4	338	4	US-09-543-681A-7759	Sequence 7759, Ap
17	53.5	17.4	713	1	US-08-190-802A-63	Sequence 63, Appl
18	53.5	17.4	713	3	US-08-477-346-63	Sequence 63, Appl
19	53.5	17.4	713	3	US-08-473-089-63	Sequence 63, Appl
20	53.5	17.4	713	4	US-08-487-072A-63	Sequence 63, Appl
21	53.5	17.4	713	4	US-09-108-857-3	Sequence 3, Appli
22	53	17.3	79	4	US-09-248-796A-26296	Sequence 26296, A
23	53	17.3	163	4	US-09-270-767-40798	Sequence 40798, A
24	53	17.3	163	4	US-09-270-767-56014	Sequence 56014, A
25	53	17.3	273	3	US-08-936-165A-395	Sequence 395, App
26	52.5	17.1	172	4	US-09-252-991A-20038	Sequence 20038, A
27	52.5	17.1	407	4	US-09-328-352-7772	Sequence 7772, Ap
28	52.5	17.1	2618	3	US-09-413-814-28	Sequence 28, Appl
29	52	16.9	559	3	US-09-134-001C-3721	Sequence 3721, Ap
30	52	16.9	574	4	US-09-902-540-11805	Sequence 11805, A
31	52	16.9	720	4	US-09-270-767-43994	Sequence 43994, A
32	52	16.9	897	4	US-09-543-681A-4915	Sequence 4915, Ap
33	51.5	16.8	119	4	US-09-270-767-41178	Sequence 41178, A
34	51.5	16.8	119	4	US-09-270-767-56394	Sequence 56394, A
35	51.5	16.8	178	4	US-09-328-352-4277	Sequence 4277, Ap
36	51.5	16.8	354	1	US-07-759-568-2	Sequence 2, Appli
37	51	16.6	114	4	US-09-489-039A-13491	Sequence 13491, A
38	51	16.6	253	4	US-09-270-767-41520	Sequence 41520, A
39	51	16.6	327	4	US-09-248-796A-16639	Sequence 16639, A
40	51	16.6	458	4	US-09-543-681A-6324	Sequence 6324, Ap
41	51	16.6	488	4	US-09-248-796A-15599	Sequence 15599, A
42	51	16.6	534	4	US-09-270-767-43905	Sequence 43905, A
43	51	16.6	815	4	US-09-489-039A-8577	Sequence 8577, Ap
44	51	16.6	873	4	US-09-543-681A-7315	Sequence 7315, Ap
45	50.5	16.4	61	4	US-09-248-796A-22227	Sequence 22227, A

ALIGNMENTS

RESULT 1
US-09-248-796A-25511
; Sequence 25511, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 25511
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-25511

Query Match 20.4%; Score 62.5; DB 4; Length 77;
Best Local Similarity 38.6%; Pred. No. 0.13;
Matches 17; Conservative 8; Mismatches 18; Indels 1; Gaps 1;

Qy 9 LFFFLQLLKQSSRRLEHT-FVFLRNFSLMLLRYIGKKRRATRFW 51
Db 11 LINFHFLLKEKEGNLENSNFFFFLDFILLLFVYLEKKKPIPRFF 54

RESULT 2
US-09-328-352-7350
; Sequence 7350, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7350
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7350

Query Match 19.7%; Score 60.5; DB 4; Length 72;

```
  Best Local Similarity   27.5%;  Pred. No. 0.23;
  Matches   11;  Conservative   17;  Mismatches    7;  Indels    5;  Gaps    2;

Qy         13 LQLLKQSSRRLEHTFVFL--RNFSLMLLRYIGKKRRATRF 50
              :| |:::| :|:| : |:  :||   :::::|:: |:  |
Db         12 MQFLRETSTQLQHAYSFMITQNF---IIKPLGQRWRSPTF 48


RESULT 3
US-09-543-681A-4564
; Sequence 4564, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
;  APPLICANT: GARY BRETON
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
;  TITLE OF INVENTION:  DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.1002-001
;  CURRENT APPLICATION NUMBER: US/09/543,681A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/128,706
;  PRIOR FILING DATE: 1999-04-09
;  NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4564
;   LENGTH: 904
;   TYPE: PRT
;   ORGANISM: Proteus mirabilis
US-09-543-681A-4564

  Query Match             19.5%;  Score 60;  DB 4;  Length 904;
  Best Local Similarity   35.3%;  Pred. No. 5.1;
  Matches   18;  Conservative    7;  Mismatches   16;  Indels   10;  Gaps    2;

Qy         15 LLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATR-------FWDPRRGTP 58
              |  |: | :||   | |  :|: || :  :| | |       | | : |||
Db        204 LAGQAERAIEH---FFRKGNLIALRELALRRMADRVDSQVKEFRDSQGGTP 251


RESULT 4
US-09-902-540-16320
; Sequence 16320, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
;  APPLICANT: Goldman, Barry S.
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Wiegand, Roger C.
;  TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
;  FILE REFERENCE: 38-10(15849)B
;  CURRENT APPLICATION NUMBER: US/09/902,540
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 60/217,883
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16320
;   LENGTH: 542
;   TYPE: PRT
;   ORGANISM: Myxococcus xanthus
US-09-902-540-16320

  Query Match             18.2%;  Score 56;  DB 4;  Length 542;
  Best Local Similarity   30.5%;  Pred. No. 11;
  Matches   18;  Conservative   10;  Mismatches   23;  Indels    8;  Gaps    2;

Qy          4 FPPSPLFFFLQLLKQSSRRL-------EHTFVFLRNFSLMLLRYIGKKRRATR-FWDPR 54
              || | :|  |   ::  : |         ||  ::|:||  |||::|      |  |: |
Db         18 FPSSEIFSTLPRRRELLKGLLGADLIGFHTVSYVRHFSGALLRHLGLDTDIDRIIWEGR 76


RESULT 5
US-09-489-039A-9391
; Sequence 9391, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
;  APPLICANT:  Gary Breton et. al
;  TITLE OF INVENTION:  NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
;  TITLE OF INVENTION:  PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE:  2709.2004001
;  CURRENT APPLICATION NUMBER: US/09/489,039A
;  CURRENT FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  US 60/117,747
;  PRIOR FILING DATE:  1999-01-29
;  NUMBER OF SEQ ID NOS:  14342
; SEQ ID NO 9391
;   LENGTH: 229
;   TYPE: PRT
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9391

  Query Match             18.1%;  Score 55.5;  DB 4;  Length 229;
  Best Local Similarity   26.4%;  Pred. No. 4.7;
  Matches   14;  Conservative    9;  Mismatches   27;  Indels    3;  Gaps    1;

Qy          4 FPPSPLFFFLQLLKQSSRRLEHTFVFLRNF---SLMLLRYIGKKRRATRFWDP 53
              | ||| :  | | :   : :   : :|      | ::| |:      | | ||
Db         51 FYPSPDYVLLNLFRDRVKAIRFLYFYLLKVLLKSTLILNYLYPTNNTTPFSDP 103


RESULT 6
US-08-403-634-4
; Sequence 4, Application US/08403634
; Patent No. 5674748
;  GENERAL INFORMATION:
;    APPLICANT:  Giordano, Antonio
;    TITLE OF INVENTION:  NOVEL HUMAN CYCLIN-DEPENDENT
;    TITLE OF INVENTION:  KINASE-LIKE PROTEINS AND METHODS
;    TITLE OF INVENTION:  OF USING THE SAME
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock, Washburn, Kurtz, Mackiewicz &
;      ADDRESSEE:  No. 5674748ris
;      STREET:  One Liberty Place, 46th floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/403,634
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/208,575
;      FILING DATE:  08-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  TJU-1482
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  316 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-403-634-4

  Query Match             18.1%;  Score 55.5;  DB 1;  Length 316;
  Best Local Similarity   35.4%;  Pred. No. 6.9;
  Matches   17;  Conservative    5;  Mismatches   23;  Indels    3;  Gaps    1;
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:24:13 ; Search time 121 Seconds
(without alignments)
245.459 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66.5	21.7	382	2	Q652R3	Q652r3 oryza sativ
2	66.5	21.7	735	2	Q704C6	Q704c6 thermoprote
3	66	21.5	482	2	Q7MV94	Q7mv94 porphyromon
4	64	20.8	524	2	Q6QHE6	Q6qhe6 brassica hi
5	63	20.5	524	2	Q6QHD9	Q6qhd9 brassica ca
6	63	20.5	524	2	Q6QHE0	Q6qhe0 brassica ol
7	63	20.5	524	2	Q6QHE1	Q6qhe1 brassica mo
8	63	20.5	524	2	Q6QHE4	Q6qhe4 brassica in
9	63	20.5	934	2	Q9HCR9	Q9hcr9 homo sapien
10	63	20.5	935	2	Q8VID6	Q8vid6 rattus norv
11	63	20.5	988	2	Q96S76	Q96s76 homo sapien
12	62.5	20.4	779	1	YU20_ARATH	Q9zpv5 arabidopsis
13	62	20.2	524	2	Q6QHE5	Q6qhe5 brassica in
14	62	20.2	524	2	Q6QHE7	Q6qhe7 brassica cr
15	62	20.2	579	2	O76734	O76734 dictyosteli
16	61	19.9	298	2	Q7X1N8	Q7x1n8 lactococcus
17	61	19.9	504	2	O13102	O13102 carassius a
18	60	19.5	511	2	Q9GHA9	Q9gha9 flagellaria
19	60	19.5	1199	2	Q9Y4J9	Q9y4j9 homo sapien
20	59	19.2	637	2	Q9W4R7	Q9w4r7 drosophila
21	59	19.2	1463	2	Q86NZ0	Q86nz0 drosophila
22	59	19.2	2243	2	Q9VGP1	Q9vgp1 drosophila
23	58.5	19.1	1976	2	Q7RIX5	Q7rix5 plasmodium
24	58	18.9	253	2	O04021	O04021 arabidopsis
25	58	18.9	374	2	Q9KCM2	Q9kcm2 bacillus ha
26	58	18.9	622	2	Q84TJ9	Q84tj9 arabidopsis
27	58	18.9	843	2	Q9SBI0	Q9sbi0 arabidopsis
28	57.5	18.7	53	2	Q8F5U4	Q8f5u4 leptospira
29	57.5	18.7	211	1	YAN8_SCHPO	Q10073 schizosacch
30	57.5	18.7	651	2	Q7D2Y4	Q7d2y4 agrobacteri
31	57.5	18.7	713	2	Q8UJI6	Q8uji6 agrobacteri
32	57.5	18.7	1082	2	Q80U35	Q80u35 mus musculu
33	57	18.6	243	2	Q8A9M4	Q8a9m4 bacteroides
34	57	18.6	492	2	Q84JR3	Q84jr3 arabidopsis
35	56.5	18.4	124	2	Q9YE31	Q9ye31 aeropyrum p
36	56.5	18.4	282	2	Q94BE8	Q94be8 kochia scop
37	56.5	18.4	332	2	O61982	O61982 caenorhabdi
38	56.5	18.4	511	1	MATK_BROIN	Q9muz3 bromus iner
39	56.5	18.4	821	2	Q73YX5	Q73yx5 mycobacteri
40	56	18.2	61	2	Q8V7E3	Q8v7e3 tt virus. o
41	56	18.2	276	2	Q8BYB4	Q8byb4 mus musculu
42	56	18.2	344	2	Q7P382	Q7p382 fusobacteri
43	56	18.2	361	2	Q9SN11	Q9sn11 arabidopsis
44	56	18.2	563	2	Q8ENR2	Q8enr2 oceanobacil
45	56	18.2	586	2	Q7WZI4	Q7wzi4 haemophilus

ALIGNMENTS

RESULT 1
Q652R3
ID Q652R3 PRELIMINARY; PRT; 382 AA.
AC Q652R3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein P0603C10.53.
GN Name=P0603C10.53;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
RT clone:P0603C10.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP005527; BAD46204.1; -.
KW Hypothetical protein.
SQ SEQUENCE 382 AA; 41577 MW; 3CADE61AE25E8DDB CRC64;

Query Match 21.7%; Score 66.5; DB 2; Length 382;
Best Local Similarity 41.5%; Pred. No. 4.7;
Matches 17; Conservative 3; Mismatches 16; Indels 5; Gaps 1;

```
Qy         18 QSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
              ::||| ||  |   |||       | :||  | | | |  |
Db        226 RASRRREHGAVLEHNFSRG-----GDRRRRRRHWPPSRRRP 261
```

RESULT 2
Q704C6
ID Q704C6 PRELIMINARY; PRT; 735 AA.
AC Q704C6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Trehalose-6-phosphate synthase-phosphatase.
GN Name=tpsp;
OS Thermoproteus tenax.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Thermoproteus.
OX NCBI_TaxID=2271;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15028704; DOI=10.1128/JB.186.7.2179-2194.2004;
RA Siebers B., Tjaden B., Michalke K., Doerr C., Ahmed H., Zaparty M.,
RA Gordon P., Sensen C.W., Zibat A., Klenk H.P., Schuster S.C.,
RA Hensel R.;
RT "Reconstruction of the central carbohydrate metabolism of
RT Thermoproteus tenax using genomic and biochemical data.";

```
RL   J. Bacteriol. 186:2179-2194(2004).
DR   EMBL; AJ621287; CAF18468.1; -.
DR   GO; GO:0003825; F:alpha,alpha-trehalose-phosphate synthase (U. . .; IEA.
DR   GO; GO:0003824; F:catalytic activity; IEA.
DR   GO; GO:0004805; F:trehalose-phosphatase activity; IEA.
DR   GO; GO:0008152; P:metabolism; IEA.
DR   GO; GO:0005992; P:trehalose biosynthesis; IEA.
DR   InterPro; IPR001830; Glyco_trans_20.
DR   InterPro; IPR006379; HAD_SF_IIB.
DR   InterPro; IPR003337; Trehalose_PPase.
DR   Pfam; PF00982; Glyco_transf_20; 1.
DR   Pfam; PF02358; Trehalose_PPase; 1.
DR   TIGRFAMs; TIGR01484; HAD-SF-IIB; 1.
DR   TIGRFAMs; TIGR00685; T6PP; 1.
SQ   SEQUENCE   735 AA;  82040 MW;  AF7A68E3844804ED CRC64;

  Query Match             21.7%;  Score 66.5;  DB 2;  Length 735;
  Best Local Similarity   37.7%;  Pred. No. 9.5;
  Matches   20;  Conservative    7;  Mismatches   11;  Indels   15;  Gaps    3;

Qy          4 FPPSPLFFFLQLLKQSSRR--LE----------HTFVFLRNFSLMLLRYIGKK 44
              |||: |   ||||    ||  ||          ||: :  |||  ::|::| |
Db        173 FPPAEL---LQLLPSEWRREILEGLLGSDLVGFHTYEYSANFSRSVVRFLGYK 222


RESULT 3
Q7MV94
ID   Q7MV94      PRELIMINARY;      PRT;   482 AA.
AC   Q7MV94;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PG1186;
OS   Porphyromonas gingivalis (Bacteroides gingivalis).
OC   Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC   Porphyromonadaceae; Porphyromonas.
OX   NCBI_TaxID=837;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=W83;
RX   MEDLINE=22829867; PubMed=12949112;
RX   DOI=10.1128/JB.185.18.5591-5601.2003;
RA   Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA   Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA   Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA   Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA   Dewhirst F.E., Fraser C.M.;
RT   "Complete genome sequence of the oral pathogenic bacterium
RT   Porphyromonas gingivalis strain W83.";
RL   J. Bacteriol. 185:5591-5601(2003).
DR   EMBL; AE017176; AAQ66284.1; -.
DR   TIGR; PG1186; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   482 AA;  53698 MW;  4E7E69ECDD3E6016 CRC64;

  Query Match             21.5%;  Score 66;  DB 2;  Length 482;
  Best Local Similarity   46.2%;  Pred. No. 7;
  Matches   12;  Conservative    5;  Mismatches    9;  Indels    0;  Gaps    0;

Qy         15 LLKQSSRRLEHTFVFLRNFSLMLLRY 40
              |||   | : | |  |||: |::|:|
Db        325 LLKSIDREVSHAFASLRNYDLIILQY 350


RESULT 4
Q6QHE6
ID   Q6QHE6      PRELIMINARY;      PRT;   524 AA.
AC   Q6QHE6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Maturase K.
GN   Name=matK;
OS   Brassica hilarionis.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=69182;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Simonetti E., Martin J.P., Gonzalez L.M., Aguinagalde I.;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY541612; AAS48145.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0008380; P:RNA splicing; IEA.
DR   InterPro; IPR008998; Agglutinin.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   Chloroplast.
SQ   SEQUENCE   524 AA;  62591 MW;  52940E1F109FCEA6 CRC64;

  Query Match             20.8%;  Score 64;  DB 2;  Length 524;
  Best Local Similarity   31.9%;  Pred. No. 14;
  Matches   22;  Conservative    9;  Mismatches    8;  Indels   30;  Gaps    4;

Qy          9 LFFFLQLLKQSSR----------------RLEHTF-VFLRNFSLML----------LRY 40
              :||||:  ||||                ::: | | ||: ||| :|          :||
Db        241 IFFFLR--KQSSHLRSTSYDVFFERILFYGKIQHFFKVFINNFSALLGLLKDPFLHYVRY 298

Qy         41 IGKKRRATR 49
              ||   ||:
Db        299 HGKYILATK 307


RESULT 5
Q6QHD9
ID   Q6QHD9      PRELIMINARY;      PRT;   524 AA.
AC   Q6QHD9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Maturase K.
GN   Name=matK;
OS   Brassica campestris (Field mustard).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3711;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Simonetti E., Martin J.P., Gonzalez L.M., Aguinagalde I.;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY541619; AAS48152.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0008380; P:RNA splicing; IEA.
DR   InterPro; IPR008998; Agglutinin.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   Chloroplast.
SQ   SEQUENCE   524 AA;  62563 MW;  00040E1F0F9FCB2A CRC64;

  Query Match             20.5%;  Score 63;  DB 2;  Length 524;
  Best Local Similarity   31.9%;  Pred. No. 19;
  Matches   22;  Conservative    9;  Mismatches    8;  Indels   30;  Gaps    4;

Qy          9 LFFFLQLLKQSSR----------------RLEHTF-VFLRNFSLML----------LRY 40
              :||||:  ||||                ::: | | ||: ||| :|          :||
Db        241 IFFFLR--KQSSHLRSTSYDVFFERILFYGKIQHFFKVFVNNFSALLGLLKDPFLHYVRY 298
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:33:28 ; Search time 51 Seconds
(without alignments)
109.423 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62.5	20.4	779	2	G84561	hypothetical prote
2	58	18.9	253	2	A86223	protein F7G19.4 [i
3	58	18.9	374	2	E83843	hypothetical prote
4	57.5	18.7	211	2	T38740	hypothetical prote
5	57.5	18.7	713	2	AC3220	catalase C [import
6	56.5	18.4	124	2	C72664	hypothetical prote
7	56.5	18.4	332	2	T33310	hypothetical prote
8	56	18.2	361	2	T45860	cyclin D3-like pro
9	56	18.2	367	2	T31750	hypothetical prote
10	56	18.2	1056	2	T00060	hypothetical prote
11	55.5	18.1	360	2	S49330	serine/threonine k
12	55.5	18.1	1019	2	T40813	probable cell divi
13	55	17.9	502	2	T41148	trp-asp repeat con
14	55	17.9	509	2	T37587	hypothetical trp-a
15	55	17.9	635	2	T27014	hypothetical prote
16	55	17.9	772	2	H84605	hypothetical prote
17	55	17.9	906	2	G70767	probable helicase
18	55	17.9	1253	2	S62544	hypothetical prote
19	54.5	17.8	155	2	AF0056	probable exported
20	54.5	17.8	643	2	E96767	unknown protein F2
21	54.5	17.8	708	2	T51026	hypothetical prote
22	54	17.6	183	2	E97888	degenerate transpo
23	54	17.6	446	2	T19625	hypothetical prote
24	54	17.6	486	2	B72554	probable Glu-tRNA
25	54	17.6	511	2	T27716	hypothetical prote
26	54	17.6	1057	2	AH2198	RND multidrug effl
27	53.5	17.4	210	2	G97235	hypothetical prote
28	53.5	17.4	705	2	B95843	probable catalase
29	53.5	17.4	713	2	JN0133	WD-40 repeat regul
30	53	17.3	160	2	D75388	NADH2 dehydrogenas
31	53	17.3	380	2	T33448	hypothetical prote
32	53	17.3	458	2	T33122	hypothetical prote
33	53	17.3	558	2	C90000	hypothetical prote
34	52.5	17.1	232	2	T49861	hypothetical prote
35	52.5	17.1	535	2	S47131	casein kinase I (E
36	52.5	17.1	650	2	S75333	oligopeptide trans
37	52.5	17.1	1085	2	JC2227	probable helicase
38	52	16.9	307	2	E90402	flagellin, probabl
39	52	16.9	343	2	S03415	hypothetical prote
40	52	16.9	399	2	B96756	hypothetical prote
41	52	16.9	498	2	D86298	hypothetical prote
42	52	16.9	570	4	B44282	retrovirus-related
43	52	16.9	633	2	T05005	hypothetical prote
44	52	16.9	1184	2	T46039	hypothetical prote
45	52	16.9	2748	2	S57976	nuclear migration

ALIGNMENTS

RESULT 1
G84561
hypothetical protein At2g18220 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84561
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84561
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-779 <STO>
A;Cross-references: UNIPROT:Q9ZPV5; GB:AE002093; NID:g4309744; PIDN:AAD15514.1; GSPDB:GN
C;Genetics:
A;Gene: At2g18220
A;Map position: 2

Query Match 20.4%; Score 62.5; DB 2; Length 779;
Best Local Similarity 34.9%; Pred. No. 5.6;
Matches 15; Conservative 9; Mismatches 16; Indels 3; Gaps 1;

```
Qy         14 QLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRG 56
              :::  : |||:|: |||  |  :| :||   : |  ||    |
Db        265 EMITFTLRRLKHSSVFLAAFPSLLRKYI---KVALHFWGTGSG 304
```

RESULT 2
A86223
protein F7G19.4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: A86223
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A86223
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-253 <STO>
A;Cross-references: UNIPROT:O04021; GB:AE005172; NID:g2342677; PIDN:AAB70399.1; GSPDB:GN

```
C;Genetics:
A;Gene: F7G19.4
A;Map position: 1

  Query Match             18.9%;  Score 58;  DB 2;  Length 253;
  Best Local Similarity   40.5%;  Pred. No. 7;
  Matches   15;  Conservative    5;  Mismatches   17;  Indels    0;  Gaps    0;

Qy         16 LKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWD 52
              |  :||||: :  |   |:|  ||:| |     | ||
Db         63 LSFASRRLDRSKSFGAMFALRGLRFIAKNDAVGRGWD 99


RESULT 3
E83843
hypothetical protein BH1549 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: E83843
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: E83843
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <STO>
A;Cross-references: UNIPROT:Q9KCM2; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BAB052
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1549

  Query Match             18.9%;  Score 58;  DB 2;  Length 374;
  Best Local Similarity   29.0%;  Pred. No. 10;
  Matches   18;  Conservative    9;  Mismatches   25;  Indels   10;  Gaps    2;

Qy          1 MQMFPPSPLFFFLQ--------LLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWD 52
              | |   : : ||:          |:    |||   :  |  || |||  :| : :   ||
Db        216 MGMQQMAAVVFFMHGIYALILFLIGSRERRL--LYFSLLTFSAMLLLLGSEEKLLSFWF 273

Qy         53 PR 54
              |:
Db        274 PQ 275


RESULT 4
T38740
hypothetical protein SPAC3H1.08c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T38740
R;Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z21808
A;Accession: T38740
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-211 <GEN>
A;Cross-references: UNIPROT:Q10073; EMBL:Z68144; PIDN:CAA92261.1; GSPDB:GN00066; SPDB:SP
A;Experimental source: strain 972h-; cosmid c3H1
C;Genetics:
A;Gene: SPDB:SPAC3H1.08c
A;Map position: 1
A;Introns: 70/2

  Query Match             18.7%;  Score 57.5;  DB 2;  Length 211;
  Best Local Similarity   48.3%;  Pred. No. 6.8;
  Matches   14;  Conservative    4;  Mismatches    6;  Indels    5;  Gaps    1;

Qy          4 FPPSPLFFFLQLLKQSSRRLEHTFVFLRN 32
              : |||     :|||||||||||     ::|
Db         32 YEPSP-----ELLKQSSRRLEQAGYSVKN 55


RESULT 5
AC3220
catalase C [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid AT
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AC3220
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AC3220
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-713 <KUR>
A;Cross-references: UNIPROT:Q8UJI6; GB:AE008687; PIDN:AAL46177.1; PID:g17743948; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: catE
A;Genome: plasmid
C;Superfamily: catalase

  Query Match             18.7%;  Score 57.5;  DB 2;  Length 713;
  Best Local Similarity   31.7%;  Pred. No. 23;
  Matches   13;  Conservative    9;  Mismatches   16;  Indels    3;  Gaps    2;

Qy         19 SSRRLEHTFVFLRNFSLMLLRYIGKKRRAT--RF-WDPRRG 56
              | |  :  :| |:  | :   |:|  | ::|  :| | |: |
Db        217 SDRTIPRSFRFMEGFGVHTFRFINAKEQSTFVKFHWKPKLG 257


RESULT 6
C72664
hypothetical protein APE0738 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: C72664
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72664
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-124 <KAW>
A;Cross-references: UNIPROT:Q9YE31; DDBJ:AP000060; NID:g5104188; PIDN:BAA79715.1; PID:d1
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0738

  Query Match             18.4%;  Score 56.5;  DB 2;  Length 124;
  Best Local Similarity   21.7%;  Pred. No. 5.4;
  Matches   18;  Conservative   10;  Mismatches   16;  Indels   39;  Gaps    3;

Qy          6 PSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRF---------------- 50
              | ||: ::  :|   |  ::          |::  |:|: || |
Db         33 PKPLYVYVHTVKSKGRIYQY---------LVVEEYLGQGRRRTILRMRLEEAVRKLLNNE 83

Qy         51 ------------WD--PRRGTP 58
                          ||  ||| ||
Db         84 KKDSAETAGWCGGWDLNPRRPTP 106


RESULT 7
T33310
hypothetical protein K02H11.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: October 16, 2005, 05:22:43 ; Search time 122 Seconds
(without alignments)
183.870 Million cell updates/sec

Title: US-10-031-158-14
Perfect score: 307
Sequence: 1 MQMFPPSPLFFFLQLLKQSS.....RYIGKKRRATRFWDPRRGTP 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	307	100.0	58	4	AAB66399	Aab66399 Human TCR
2	307	100.0	58	7	ADB75581	Adb75581 Prostate
3	307	100.0	58	8	ADP54863	Adp54863 Human PRO
4	91	29.6	16	4	AAB67199	Aab67199 Human TCR
5	63	20.5	934	4	AAG62677	Aag62677 Human typ
6	63	20.5	935	4	AAG62680	Aag62680 Rat type
7	62.5	20.4	688	3	AAG32065	Aag32065 Arabidops
8	62.5	20.4	688	3	AAG40595	Aag40595 Arabidops
9	62.5	20.4	728	3	AAG32064	Aag32064 Arabidops
10	62.5	20.4	728	3	AAG40594	Aag40594 Arabidops
11	62.5	20.4	764	3	AAG40593	Aag40593 Arabidops
12	62.5	20.4	779	3	AAG32063	Aag32063 Arabidops
13	62	20.2	16	4	AAB67200	Aab67200 D discoid
14	61	19.9	116	4	AAO06173	Aao06173 Human pol
15	60.5	19.7	72	6	ADA36063	Ada36063 Acinetoba
16	60.5	19.7	1855	2	AAY21803	Aay21803 B. subtil
17	60.5	19.7	1855	3	AAY83271	Aay83271 Polypepti
18	60	19.5	85	4	AAO07716	Aao07716 Human pol
19	60	19.5	900	6	ABU40841	Abu40841 Protein e
20	60	19.5	904	7	ADF04279	Adf04279 Bacterial
21	59	19.2	641	4	ABB59719	Abb59719 Drosophil
22	59	19.2	2176	4	ABB61386	Abb61386 Drosophil
23	58	18.9	83	4	AAO09340	Aao09340 Human pol
24	57.5	18.7	135	5	ADK34310	Adk34310 Novel hum
25	57.5	18.7	319	8	ABM82552	Abm82552 Human dia
26	57.5	18.7	673	8	ADS25657	Ads25657 Bacterial
27	57.5	18.7	673	8	ADS26140	Ads26140 Bacterial
28	57.5	18.7	703	8	ADS22475	Ads22475 Bacterial
29	57	18.6	134	4	AAO01316	Aao01316 Human pol
30	56.5	18.4	119	4	AAO03520	Aao03520 Human pol
31	56	18.2	78	4	AAG74605	Aag74605 Human col
32	56	18.2	232	3	AAG07077	Aag07077 Arabidops
33	56	18.2	232	3	AAG53866	Aag53866 Arabidops
34	56	18.2	308	3	AAG07076	Aag07076 Arabidops
35	56	18.2	308	3	AAG53865	Aag53865 Arabidops
36	56	18.2	361	3	AAG53864	Aag53864 Arabidops
37	56	18.2	361	3	AAG07075	Aag07075 Arabidops
38	56	18.2	459	8	ADS30236	Ads30236 Bacterial
39	55.5	18.1	229	7	ABO62874	Abo62874 Klebsiell
40	55.5	18.1	313	8	ADR50796	Adr50796 Human c-b
41	55.5	18.1	314	7	ADI63090	Adi63090 Human apo
42	55.5	18.1	314	8	ADO57044	Ado57044 Human cyc
43	55.5	18.1	316	2	AAW04872	Aaw04872 Phosphory
44	55.5	18.1	316	2	AAW04870	Aaw04870 Cyclin-de
45	55.5	18.1	323	8	ADF88822	Adf88822 Human kin

ALIGNMENTS

RESULT 1
AAB66399
ID AAB66399 standard; protein; 58 AA.
XX
AC AAB66399;
XX
DT 12-APR-2001 (first entry)
XX
DE Human TCRgamma alternate reading frame protein.
XX
KW Human; TARP; prostate cancer; breast cancer; immunotherapy;
KW T cell receptor gamma alternate reading frame protein; TCRgamma.
XX
OS Homo sapiens.
XX
PN WO200104309-A1.
XX
PD 18-JAN-2001.
XX
PF 12-JUL-2000; 2000WO-US019039.
XX
PR 13-JUL-1999; 99US-0143560P.
PR 01-OCT-1999; 99US-0157471P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Pastan I, Essand M, Lee B, Vasmatzis G, Wolfgang C;
XX
DR WPI; 2001-081050/09.
DR N-PSDB; AAF56401.
XX
PT Isolated T-cell receptor gamma alternate reading frame protein useful for
PT diagnosing and raising an immune response to prostate cancer and breast
PT cancer.
XX
PS Claim 2; Fig 1; 85pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC human T cell receptor alternate reading frame protein (TARP). This
CC protein is expressed in prostate and breast cancer cells at higher levels
CC than normal and so can be used in the immunotherapy of these cancers, as
CC well as their detection and prevention
XX
SQ Sequence 58 AA;

Query Match 100.0%; Score 307; DB 4; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.6e-36;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
```

RESULT 2
ADB75581
ID ADB75581 standard; protein; 58 AA.
XX
AC ADB75581;
XX
DT 04-DEC-2003 (first entry)
XX
DE Prostate cancer marker protein.
XX
KW Prostate; cancer; cytostatic; gene therapy; marker.
XX
OS Homo sapiens.
XX
PN WO2003009814-A2.
XX
PD 06-FEB-2003.
XX
PF 25-JUL-2002; 2002WO-US023913.
XX
PR 25-JUL-2001; 2001US-0307982P.
PR 22-AUG-2001; 2001US-0314356P.
PR 25-SEP-2001; 2001US-0325020P.
PR 12-DEC-2001; 2001US-0341746P.
PR 05-MAR-2002; 2002US-0362158P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbatcheva B;
PI Hoersh S, Kamatkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
XX
DR WPI; 2003-248033/24.
XX
PT New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX
PS Disclosure; SEQ ID NO 405; 99pp; English.
XX
CC The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 58 AA;

```
  Query Match             100.0%;  Score 307;  DB 7;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-36;
  Matches   58;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQMFPPSPLFFFLQLLKQSSRRLEHTFVFLRNFSLMLLRYIGKKRRATRFWDPRRGTP 58
```

RESULT 3
ADP54863
ID ADP54863 standard; protein; 58 AA.
XX
AC ADP54863;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human PRO protein sequence SEQ ID NO:839.
XX
KW human; PRO; immune related disease; inflammatory immune response;
KW immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW virucide; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2004039956-A2.
XX
PD 13-MAY-2004.
XX
PF 28-OCT-2003; 2003WO-US034381.
XX
PR 29-OCT-2002; 2002US-0422472P.
XX
PA (GETH) GENENTECH INC.
XX
PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
DR WPI; 2004-376182/35.
DR N-PSDB; ADP54862.
XX
PT New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT and treating an immune related disease, e.g. systemic lupus
PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT stimulating an immune response.
XX
PS Claim 1; SEQ ID NO 839; 3009pp; English.
XX
CC The present invention describes an isolated PRO nucleic acid (I). Also
CC described: (1) a vector comprising (I); (2) a host cell comprising the
CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC antibody which specifically binds to a polypeptide of (4); (7) a
CC composition of matter comprising a polypeptide of (4), an agonist or
CC antagonist of the polypeptide or an antibody that binds to the
CC polypeptide in combination with a carrier; (8) an article of manufacture
CC comprising a container, a label on the container and a composition of
CC matter of (7); (9) a method of treating an immune related disease in a
CC mammal; (10) a method for determining the presence of a PRO polypeptide
CC in a sample suspected of having the polypeptide; (11) a method of
CC diagnosing an immune related disease or an inflammatory immune response
CC in mammal; (12) a method of identifying a compound that inhibits or
CC mimics the activity of or expression of a gene encoding a PRO polypeptide
CC ; and (13) a method of stimulating the immune response in a mammal. The
CC PRO sequences have antiallergic, antianaemic, antiarthritic,
CC antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC virucide activities, and can be used in gene therapy. The nucleic acid
CC (I) and the encoded polypeptides, compositions, kits and methods are
CC useful in diagnosing and treating an immune related disease and in
CC stimulating an immune response. The present sequence represents a human
CC PRO protein from the present invention.
XX
SQ Sequence 58 AA;

```
  Query Match             100.0%;  Score 307;  DB 8;  Length 58;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-36;
```